The Thesis Committee for Rebekah Elizabeth Post
Certifies that this is the approved version of the following thesis:

Code-switching in the Determiner Phrase:

A comparison of Tunisian Arabic-French and Moroccan Arabic-French switching

APPROVED BY

SUPERVISING COMMITTEE:

Supervisors:

Barbara E. Bullock

David P. Birdsong

Code-switching in the Determiner Phrase:

A comparison of Tunisian Arabic-French and Moroccan Arabic-French switching

by

Rebekah Elizabeth Post, BA

Thesis

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

Master of Arts

The University of Texas at Austin

December 2010

Acknowledgements

Thanks goes first to my thesis advisors, Barbara Bullock and David Birdsong, without whom I wouldn't have even known where to begin and whose thoughts, advice, suggestions, guidance and time has been invaluable over the preparation of this thesis. I am also indebted to the native speakers of Tunisian and Moroccan Arabic, both in and out of Austin, that generously gave their time and knowledge to help me better understand their dialects and create the survey that served as the focus of this study. This study could not have been completed without the assistance of all the Moroccans and Tunisians who participated in the survey. Finally, to all of my friends and family who encouraged me, worked with me, and gave me much-needed perspective and distractions throughout the entire process. I could not have done it without any of you.

December 3, 2010

Abstract

Code-switching in the Determiner Phrase:

A comparison of Tunisian Arabic-French and Moroccan Arabic-French switching

Rebekah Elizabeth Post, MA

The University of Texas at Austin, 2010

Supervisor: Barbara E. Bullock, David P. Birdsong

Code-switching (CS) between French and Arabic is common across North Africa and in parts of the Middle East. Many researchers have examined this phenomenon in Tunisia (Belazi 1991, Lawson & Sachdev 2000, Belazi et. al 1994) and Morocco (Abbassi 1977, Bentahila 1983, Bentahila & Davies 1983, Lahlou 1991, Redouane 2005.) Corpus and elicited data from these two countries has helped form the basis of proposed universal constraints on code-switching, specifically the Functional Head Constraint (FHC) (Belazi et al 1994) and the Complement Adjunct Distinction (CAD) (Mahootian and Santorini 1996). However, CS between French and Moroccan and Tunsian dialects has not been directly investigated within a single study. This study is a step in filling that gap.

Using a web-based survey, the present study examines native dialect speakers' ratings of authenticity of sentences that contain both French and Arabic with a switch occurring in the Determiner Phrase (DP). The syntactic structure of the DP in the dialects examined is the same,

(DP = D (D) N (A)). This is similar to the DP in French (DP = D (A) N (A)) with a few key differences that make it possible to test the FHC and CAD within the DP alone. An example of one of the eight possible switch types, between an Arabic Demonstrative Determiner and a French Definite Determiner, is seen here between Moroccan Arabic and French: *Men dima had* l'homme n'aime pas les chiens. (*Since always this* the man doesn't like dogs.) A mixed-model ANOVA performed on the participants' ratings reveals main effects for dialect, sex and switch type. Significant interactions also exist, including an interaction between switch type, sex and dialect.

While further research is needed, the results indicate that syntactic constraints may not be the only way to understand the practice of CS. Instead, a typological approach, as suggested by Muysken (2000), may lead to a more complete understanding of why and how communities use multiple languages.

Table of Contents

List of Tables

List of Figures

1. Introduction

The focus of the current study is code-switching (CS) between French and Arabic. The two languages have long been in contact in North Africa, among other regions, leading to the frequent occurrence of French-Arabic code-switching, particularly in the Maghreb. The language pair has been of particular interest due to the fact that they are typologically distinct languages. In many cases, the syntax of the two languages do not align, making attested switches between the two difficult to account for in many proposed models of code-switching.

While French and Arabic differ syntactically in many respects, the present focus is on the Determiner Phrase (DP). Two possible DPs in each language are seen below, in Arabic (1) and in French (2).

(1) (a) l-medina
"the city"

(b) had l-medina
"this city"

(2) (a) la ville
"the city"

(b) cette ville
"this city"

Constraints on CS based on linear equivalence, such as Poplack's Equivalence Constraint, focus on the surface word order. Equivalence is found in French and Arabic in the DP when the Determiner (DET) is a Definite Determiner (DEF), seen in (3a, b).

(3) (a) $_{DP}$[l-$_{NP}$[medina]]

(b) $_{DP}$[la $_{NP}$[ville]]

However, when the DET is a Demonstrative Determiner (DEM), the Arabic dialects discussed here require a DP beginning with a DEF to follow it, as in (4a) where French requires a noun.

(4) (a) $_{DP}$[had $_{DP}$[l-$_{NP}$[medina]]]

(b) $_{DP}$[cette $_{NP}$[ville]]

As the demonstratives in each language are followed by a different complement, constraints on

CS based on linear equivalence do not allow switches involving French and Arabic to occur between a DEM and a noun. This type of switch does occur, as attested in Moroccan Arabic-French code-switching:

(5) dak *la chemise*
that *the shirt*

(6) *cette* xubza
this *loaf*
(=69, 92 in Bentahila & Davies 1983)

Several syntactic constraints on code-switching have been proposed that take into consideration French-Arabic data. These models will be examined to reveal the similarities and differences between them. As will be seen, the models diverge regarding code-switching within the DP. For this reason, in this study code-switching in the DP is explored experimentally in order to identify which, if any, proposed constraints are consistent with actual speakers' preferences.

Section 2 contains the background and context relevant to the study, including a review of the literature on syntactic constraints on code-switching, an explanation of the determiner phrases under consideration and the sociolinguistic factors relevant to them. Section 3 is a description of the experimental procedure of the study and a presentation of the results. Section 4 is a discussion of the results of the study.

2. Background

In order to situate this study relative to the existing literature on the phenomenon of code-switching, previous studies on code-switching are discussed in section 2.1. As will be seen in 2.1, code-switching must be understood in relation to the languages involved and factors that influence the use of those languages within the communities that speak them. A description of Arabic and French language use in Morocco and Tunisia is given in section 2.2 and 2.3 presents the relevant sociolinguistic background.

2.1 Background on code-switching and syntax

2.1.1 Definition of code-switching for this study

The term Code-switching can be used to describe a variety of uses of multiple languages. Timm (1975) defined it as "that preeminently bilingual mode of communication characterized by frequent shifts from one language to the other (typically without phonological interference) throughout the flow of natural conversation" (473.) Code-switching has also been the term used to refer to a bilingual speaker's ability to choose one language over another for various pragmatic reasons (Kachru 1977). Poplack defined it as "the alternation of two languages within a discourse, sentence or constituent" (1980, 581). Her definition expands on what one might assume from Timm's; any use of two languages, whether in conversation or another speech setting, qualifies as code-switching. Bentahila and Davies (1983) return to a definition close to that of Timm's, "the use of two languages within a single conversation, exchange or utterance," (302) but distinguish it from borrowing, which they define as "the use in one language of items which originate in another language, but which are currently felt to form an integrated part of the borrowing language" (302). The difference for them seems to be

the acceptance of the borrowed word by speakers of the 'borrowing language.' Determining how speakers of the language feel about a specific word may be difficult as it likely varies between speakers. For this reason determining whether a given word constitutes a borrowing or a single word insertion is problematic.

Myers-Scotton (1993) gives a more specific definition of code-switching that is specific to the model she proposes, "the selection by bilinguals or multilinguals of forms from an embedded variety (or varieties) in utterances of a matrix language during the same conversation" (3). Later, she refines this to be 'classic CS,' defined as "CS in which empirical evidence shows that abstract grammatical structure within a clause comes from only one of the participating languages" (2009, 337). Muysken (2000) chooses to distinguish between code-mixing, code-switching and lexical borrowing. Code-mixing, as defined by Muysken, is similar to Timm's code-switching. That is, it is restricted to a single sentence, defined as "all cases where lexical items and grammatical features from two languages appear in one sentence" (1). He reserves code-switching more specifically "for the rapid succession of several languages in a single speech event" (1). Within Muysken's typology, lexical borrowing is considered a specific type of insertion.

A common difficulty in the literature is determining whether borrowing is a distinct process from code-switching. Pfaff (1979) gives a summary of definitions of borrowing. She considers that borrowing is a process that monolinguals may participate in while code-switching is found only in the speech of bilinguals. Some definitions of borrowing are based on the morpho-syntax of the borrowed word, while others refer to the speaker's opinion on whether the borrowed word is a part of the primary language used, or if an equivalent to the word exists at all in the primary language. Poplack's defines borrowing as a lexical item that is structurally or

phonologically integrated into the borrowing language, whereas code-switches are not (1980). Definitions that appeal to the lexicon of the languages involved classify a lexical item as a borrowing if an item in the primary language does not exist with same nuances of meaning as the borrowed item. This can make it easier to classify any word as a borrowing as it is unlikely that all of the nuances in meaning of a word from one language will map exactly onto a single lexical item from another. The difficulty of delineating between borrowing and code-switching persists today (see Bullock and Toribio 2009).

The focus of this study is intrasentential code-switching rather than borrowing. The use of borrowed words, here using Bentahila and Davies' definition of words considered part of the borrowing language by speakers, has been avoided where possible as described in section 3.2.1. All experimental stimuli follow the pattern of alternational, rather than insertional, code-switching as described by Muysken (2000). No statistics on this frequency of this type of code-switching between French and Arabic exist. One study suggests that single-word switches are the most common in Tunisian Arabic-French (Sayahi 2007) while others indicate that a wide variety of switches is possible in Moroccan Arabic-French (Bentahila and Davies 1983, 1995). The alternational pattern is used here to create a controlled environment for closer examination of switch locations.

2.1.2. Summary of existing literature

Code-switching has sparked the interest of linguists since the early 20th century (Espinosa 1911) with formal investigation of it beginning in the 1970s. The early studies viewed code-switching as a random process, without patterns, and a sign of incomplete language learning (Weinreich 1953), but by the mid-1970s researchers found that in fact clear patterns

can be observed in code-switching by a specific speaker and within a community (Gumperz and Hernández-Chávez1971, Blom and Gumperz 1972).

With the rise of generative grammar, interest in the phenomenon of CS for its implication on linguistic theory turned to the possible syntactic constraints on code-switching. The earliest constraints were based on surface word order, the most influential example of which is Poplack's Two Constraint Model (1980), discussed in greater detail in section 2.1.4 below. An approach based on surface word order seems to accurately account for the code-switching in some language pairs, but far from all. For this reason, generative grammar became a popular framework for exploring syntactic constraints on CS (DiSciullo et al. 1986, Bentahila and Davies 1983, Belazi et al. 1994, Mahootian 1993, Mahootian and Santorini 1996, MacSwan 2009). While many linguists have participated in the search for universal syntactic constraints on code-switching, not all agree that these exist. Others have suggested that any constraints must be language-specific (Clyne 1987.) Related to this claim is Bhatt's unique model based on Optimality theory, stating that constraints may apply to all language pairs, but may be ranked differently by each.

Researchers have also proposed models not based on generative grammar, the most influential of which is Myers-Scotton's Matrix Language Frame, which focuses on the asymmetrical use of languages involved in code-switching, described in 2.1.4.4 below. Muysken (2000) takes a different approach instead identifying three code-switching patterns, each associated with a specific social context, as seen in 2.1.4.5. He states that a community may use one, two or all three patterns of code-switching. Other aspects of code-switching, such as the sociolinguistic and pragmatic variables that affect its use, have been the topic of many other studies, but syntactic constraints have been generally considered apart from these other

variables, with the notable exception of Muysken. If Muysken is correct that sociolinguistic and pragmatic factors are critical in code-switching, studying them separately from syntactic constraints may have hindered the full exploration of the subfield. Various language pairs have been examined in studies on code-switching, but there seem to be few thus far that directly compare switching between one language and a second language used in multiple contexts, or between different dialects of the same language (Poplack 1987). If unique switching patterns are found in separate communities using the same language, then any comprehensive model should be able to account for this variation.

2.1.2.1 Literature on Arabic-French code-switching

The bulk of literature on code-switching between French and a dialect of Arabic focuses on Moroccan Arabic (Abbassi 1977, Bentahila 1983, Bentahila& Davies 1983, Bentahila& Davies 1995, Lahlou 1991, Redouane 2005) and Tunisian Arabic (Belazi 1991; Lawson &Sachdev 2000; Belazi etal. 1994) with little to no research on switching between French and other dialects of Arabic. Within this literature, while there have been some studies exploring the syntactic constraints of switching (Bentahila and Davies 1983, Mahootian &Santorini 1996, Belazi et al. 1994) the emphasis has been on the language attitudes regarding it (Bentahila 1983a, Bentahila 1983b, Lawson & Sachdev 2000). As will be seen below, the authors of two influential studies on CS constraints that were based in part on data from French-Arabic code-switching come to significantly different conclusions. Both sets of authors state that their proposed constraints are universal. The fact that they are based on the same language pair with data from different countries is not acknowledged as a possible source of divergence. It is possible that the apparent contradictions between the theories are due to differing usage in the communities

from which the data came. Mahootian and Santorini (1996) based their analysis on Bentahila and Davies's (1983) Moroccan Arabic-French data, while Belazi et al. (1994) used Tunisian Arabic-French data.

An alternative to universal constraints is that each language pair, or separate communities sharing a language pair, may prefer different switch types. This is assumed by Bhatt (1997) in his Optimality Theoretic account of code-switching. It is also the position taken by Muysken (2000.) As the social settings for dialects of Arabic can vary considerably, it is possible that different dialects prefer different kinds of constraints on code-switching for either grammatical or sociolinguistic reasons; both must be considered in order to reach a valid conclusion.

2.1.3 Generative approaches

Universal constraints are considered the goal by many linguists who maintain that there are syntactic rules that govern speech in all languages. This notion comes from Chomsky's proposition that all humans have a language-learning mechanism called universal grammar. In this theory, all languages share common categories and rules. Any time a language diverges from the expectations laid out in universal grammar, it is considered an exception that must be explained. This idea is central to generative grammar, which is the framework on which both the Belazi et al. and the Mahootian and Santorini studies are based. Within generative grammar, variation, including differences between code-switching in different language pairs, is often ignored or down-played when it does not support universal grammar (Featherston 2007, Evans and Levinson 2009). The goal here is not to propose a new constraint or to refute an existing model, but to investigate the possibility that dialectal differences may be the source of

contradictory results in previous studies on Arabic-French code-switching.

The earliest models of code-switching were based on surface word order, the Equivalence Constraint mentioned above is an example of this. As is discussed in section 2.1.4, linear order fails to account for many possible switches or to rule out impossible switches. For this reason, the hierarchical approach to syntax of Generative Grammar has often been the preferred framework used to propose syntactic constraints on code-switching since it is not strictly based on linear order. An overview of the relevant aspects of this framework is presented here.

The main premise of Generative Grammar is that it is possible to identify the rules of a language that both describe and predict the structure of a grammatical sentence, including word order and presence and/or absence of appropriate elements. Within this theory, all utterances are made up of phrases. A phrase is a constituent of a sentence whose syntactic properties are determined by the head; the structure of a phrase is determined by the head-complement relationship. For instance, a Determiner Phrase (DP) is headed by a Determiner, which selects a nominal element as its complement. A head selects for a complement according to selectional restrictions, or restrictions on the types of lexical elements that may be its complements and adjuncts. Two common types of restrictions are c-selection and s-selection. C-selection refers to the syntactic category that a head selects and s-selection to the semantic information regarding what a head may select, described in terms of theta-roles, which are semantic roles that describe the role of a noun phrase relative to a verb such as agent and patient. Abney (1987) proposes an additional selection mechanism, f-selection, which is the way that a functional (as opposed to lexical) head selects its complement. In generative grammar, a complement is governed by its head while an adjunct is not.

In the French sentence, “Les petits enfants aiment la glace,” “Les petits enfants” is the Determiner Phrase (DP). For the French DP, and the NP that is its complement, a simple form of the phrase structure rules can be seen in (7).

(7) DP → D NP
NP → (AP) N (AP) (PP)

The status of the Determiner Phrase has been much debated. Abney (1987) first suggested that the Determiner heads the phrase, arguing against the original generative view that proposed that the Noun is the head of the phrase (NP) while the Determiner is an adjunct of the noun phrase. Haegeman (1991) notes that Abney’s proposal is controversial. The proposal of the DP as a functional projection over the Noun is parallel to the Complementizer Phrase (CP) where the complementizer is the head of a Verb. The acceptability of sentences with a bare noun in languages such as English, as in (8), may indicate that the Determiner is an optional part of the Noun Phrase.

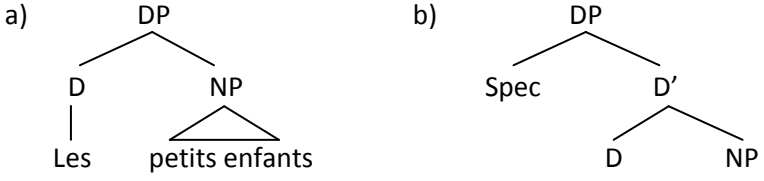
(8) Cats chase mice.

In other languages, such as French, the Noun is an optional element, lending weight to the argument for a Determiner Phrase, as in (9)

(9) Les petits aiment la glace.
The littles like the ice cream.
‘The little ones like ice cream.’

The structure tree of a basic French DP is seen in figure 1a, while 1b shows the structure of the

Figure 1 – The Determiner Phrase in French



phrase without any lexical insertion. Abney finds morphological, syntactic and semantic evidence for the DP. Morphological motivation for the DP is found in languages that show identical agreement on subject and verb. Syntactic motivation is found in argument structure because nouns can take both internal and external arguments in the same way as verbs. Relative word order within the DP, specifically of an adjective in respect to a noun, again shows that a noun moves in ways similar to a verb within a clause; languages that have postverbal adverbs tend to have postnominal adjectives while those with preverbal adverbs often have prenominal adjectives. This movement can be accounted for by a assuming a DP, but not an NP, analysis. Semantic evidence for the DP follows the fact that an NP is a non-argument while a DP is an argument.

Bernstein (2001) illustrates the existence of a DP in Romance languages specifically, citing the word order of adjectives and nouns in Romance languages. In some Romance dialects, such as Walloon, adjectives are strictly pre-nominal. In French many are pre-nominal, but a large number are post-nominal. In Sardinian, a dialect of Italian, adjectives are always post-nominal. Bernstein states that the position of the adjective relative to the noun is due to the noun raising to different positions within the DP. Additional support of the DP hypothesis is found in the co-occurrence of a prenominal demonstrative and a postnominal reinforcer in Romance languages, including French. An example of this is in (10).

(10) Ce livre-ci
This book-here
"This book" (=26 in Bernstein 2001, 552)

Bernstein proposes that both the demonstrative and deictic marker are generated to the left of the noun as a head and specifier of a functional projection above the noun. The demonstrative then raises to the D position and the noun raises, crossing over the reinforcer, to its surface

position between the two. In this study, it is assumed, following Bernstein (2001), Abney (1987) and many others, that the DP is the correct analysis for French.

2.1.4 Code-switching models

While the social and pragmatic reasons for code-switching cannot be ignored, the focus here is on the syntactic constraints involved, particularly those that have been proposed to account for CS between Arabic and French. All constraints will be examined in reference to their implications within the determiner phrase.

Grammatical constraints on code-switching were proposed from the mid-70s (Gingras 1974, Timm 1975, Pfaff 1979). However, the first model of code-switching to gain wide-spread attention was put forth by Poplack in 1980. Examining data from Spanish-English code-switching, she identified two constraints that she proposed were universal: the Free Morpheme Constraint and the Equivalence Constraint. The Free Morpheme Constraint states that:

> (11) A switch may not occur between a bound morpheme and a lexical item unless the latter has been phonologically integrated into the language of the bound morpheme. (Sankoff and Poplack 1981: 5)

The classic example of this is the hypothetical form 'eat-*iendo*.' This represents a switch between the English word 'eat,' a free morpheme, and the Spanish present progressive suffix, 'iendo.' (Poplack 1980: 586) As *eat* does not follow Spanish phonology, Poplack asserts that the switch cannot occur based on her corpus. The Equivalence Constraint was first formulated by Poplack as a tendency (1980:586) but was soon refined to be a strict rule as follows:

> (12) The order of sentence constituents immediately adjacent to and on both sides of the switch point must be grammatical with respect to both languages involved simultaneously. This requires some specification: the local co-grammaticality or

> equivalence of the two languages in the vicinity of the switch holds as long as the order of any two sentence elements, one before the switch point and one after the switch point, is not excluded in either language. (Sankoff and Poplack 1981: 5-6)

This works relatively well for a pair of languages that often have similar surface structure, as do English and Spanish. The classic example is seen in (13).

(13)(a) I told him that so that he would bring it fast.

(b) *(Yo) le deje eso pa'que (él) la trajera ligero.*

(c) I told him that *pa'que la trajera ligero.* (Poplack 1980: 586)

In (13a) the proper structure of an English sentence is seen, followed by the same sentence in Spanish (13b). In (13c) a code-switch occurs between the two between a pronoun (that/*eso*) and a complementizer (so that/*pa'que*), illustrating the Equivalence Constraint. However, other languages do not share such structural similarities, but CS between them does happen. One syntactic location where non-equivalence may occur is between an adjective and a noun where the languages involved differ with regard to adjective placement. Di Sciullo et al. illustrate this with an Italian-English example, seen here in (14).

(14)Ma cistanno dei *smart* Italiani
but there are of-the smart Italian
'But there are smart Italians' (=40a in DiSciullo et al. 1986, 155)

In Italian, the adjective would be found to the right of the noun it modifies; in this sentence it is found to the left, its position in English. Since Poplack's constraints were first proposed, numerous examples from many language pairs have shown that they are inadequate for describing observed language use.

Data from Arabic-French CS has also been used to refute the Equivalence Constraint. In addition to the examples given above in section 1, Belazi (1991) shows that even where there is similar word order in French and Arabic, switches may not be permitted, as in (15).

(15) *On se voit à la *zaeːm9a*
We see each other at the university
'We will meet at the university.' (=33 Belazi 1991, 208)

The code-switch in (15) does not violate the Equivalence Constraint as the word order is the same in both languages, but is not accepted by native speakers of Tunisian Arabic. In other instances, the surface order differs and yet the switch is acceptable, as in (16).

(16) *Wahed* le liquide
one the liquid
'one liquid' (=70 Bentahila and Davies 1983, 317)

In 16 a switch is made in the determiner phrase. In Arabic, the number '*wahed'* must be followed by a definite article, although in French its equivalent, '*un*' must be followed directly by a lexical noun. The Equivalence Constraint thus fails for French-Arabic switching on two grounds; it is both not restrictive enough and too restrictive. It allows switches that match in surface structure but are not acceptable to bilingual speakers, and does not allow code-switches where languages do not align but for which examples are attested.

Two major constraints on CS in general are based in part on data from French-Arabic code-switching. The first is the Functional Head Constraint, proposed by Belazi, Rubin and Toribio (1994). The other constraint is the Complement/Adjunct Distinction, as proposed by Mahootian and Santorini (1996). While each of these constraints was originally presented as universal, Belazi et al. (1994) based their findings on switches between Tunisian Arabic and French as well as Spanish and English. Mahootian and Santorini (1996) found contradictory evidence for Belazi et al.'s constraint through switches between Moroccan Arabic and French, attested by Bentahila and Davies (1983). Mahootian and Santorini went a step further by also including a large amount of data from other languages to support the generalizability of the

model that they propose. A third model was proposed by Chan (2003) that uses the parts of these constraints that seem accurate, while modifying them to better explain the observed data. This model is called the Functional Head Selection Constraint.

2.1.4.1 The Functional Head Constraint

Drawing on feature checking, introduced by Chomsky (1993) as part of the Minimalist Program, and Abney's (1987) theory of 'f-selection,' Belazi et al. (1994) proposed the Functional Head Constraint (FHC). Based on the framework of generative grammar and X-Bar theory, this constraint relies on the notions of functional versus lexical categories, heads, complements and adjuncts. It is defined in Belazi et al. (1994) as in (17)

(17) *The language feature of the complement f-selected by a functional head, like all other relevant features, must match the corresponding feature of that functional head.*

= (16) in Belazi et al. (1994)

Belazi et al. posit that a language feature is a marked on all words in all languages. In monolingual speech the language feature of a complement matches that of its functional head by default. It is then only in code-switching contexts that the language feature can be observed. The functional heads listed by Belazi et al. are Complementizer, Inflection, Determiner, Quantifiers, and Negation. Lexical heads do not constrain switching according to the FHC.

Within the DP, the FHC holds that a switch between the Determiner and its complement or between the Quantifier and its complement is not permitted. This would rule out a switch between a French determiner and a Moroccan Arabic noun, an example of which is given in (18).

(18) *Cette *bint*
This *girl*
'This girl'

This type of switch is not possible according to the FHC because DET, a functional head, is the head of the DP and N is its complement. If DET is in French, it carries the language feature [French] while an Arabic noun carries the language feature [Arabic]. Given that the language features do not match, the switch is deemed impossible. A switch between a lexical head, here a noun, and its adjunct, an adjective, as in (19) is permitted because the language features do not need to match.

(19) Cette fille *zwina*
This girl *pretty*
'This pretty girl'

Switches between a functional head and its complement were not found in Belazi's data (1991). Sequences of an Arabic DET and a French N do occur, but are considered borrowings. Lexical items were classified as borrowings when they filled a lexical gap in the primary language. No instances of a French DET followed by an Arabic N were attested; Belazi maintains that if this were a type of switch, and not borrowing, then the process would happen in both directions. As the nouns in the attested sequences of an Arabic DET followed by a French N are borrowings, not switches, the FHC is not violated by them.

Mahootian and Santorini (1996) found many attested code-switches that contradict this constraint in the Moroccan Arabic-French switching reported by Bentahila and Davies (1983) as well as in other language pairs. Two of these are seen in (20) and (21).

(20) lorsque j'ai vue que *mabqaš*
'when I saw that *there was nothing left*'
(French-Moroccan Arabic, =22 in Bentahila and Davies 1983)

(21) Oui, alors j'ai dit que *si potev* aller comme ça.
'Yes, so I said that *we could* go like that'
(French-Italian, =37b in Di Sciullo et al. 1986)

The examples in (20) and (21) show switches between a complementizer, *que* in both cases, and

its complement. (21) also contains a switch between the inflected verb *potev* and its complement. As complementizers and inflection are considered functional heads, the FHC would prohibit these switches. While these counter-examples indicate that the FHC is not universal, as discussed in 2.1.4.2, it is clear that the definition of borrowing used would greatly affect the type of instances of code-switching found within a corpus.

Without any additional restrictions on code-switching, the FHC would allow a switch between a noun in French and an adjective that either precedes or follows it in Arabic because French adjectives may be found on either side of the noun, as in (22).

(22) a. La jolie boîte
The pretty box
b. La boîte noire
The box black
'The black box'

A small number of high-frequency adjectives, including 'jolie' in (22a) are precede the noun they modify while most, including colors, follow it as in (22b). On observing that switches do not occur completely freely between lexical heads and their complements and adjuncts, the Word Grammar Integrity Corollary was also proposed by Belazi et al. (1994), as seen in (23)

(23) *A word of language X, with grammar G_x, must obey grammar G_x.*
(=28 in Belazi et al. 1994)

This constraint applies to code-switching in all syntactic environments and applies only to the syntax, not the language of the involved elements. Unlike the Equivalence Constraint cited above, the Word Grammar Integrity Corollary is based on hierarchical syntax. However, Belazi et al.'s formulation is subject to the same problem regarding word order of nouns and adjectives as noted for the Equivalence Constraint. Specifically, it would prohibit a switched pre-nominal adjective before a noun in a language that only allows post-nominal adjectives, as in (14) above.

For this reason, the Word Grammar Integrity Corollary has not been as widely accepted as the Functional Head Constraint.

Belazi et al.'s data is based on elicited grammaticality judgments that were later compared to natural conversation data. Others see the use of grammaticality judgments as a weakness (Mahootian & Santorini 1996) as Rickford's (1975) work indicates that judgments may indicate stereotypes instead of reflecting the true usage of the speaker or community. Despite this potential drawback, elicited acceptability judgments remain a commonly-used way to systematically explore whether theoretically possible syntactic structures are acceptable to native speakers, even if rarely used. The primacy of natural data cannot be understated, but well-crafted judgment elicitation may highlight less commonly employed, but nevertheless acceptable, types of code-switching. The utility of acceptability judgments is further discussed in section 3.2.

2.1.4.2 The Complement/Adjunct Distinction

Mahootian and Santorini (1996) proposed the Complement/Adjunct Distinction (CAD) as a response to the Functional Head Constraint. They do not agree that a language feature as proposed by Belazi et al. is an appropriate analysis, stating that its existence is visible only in code-switching and therefore poorly motivated. They instead rely only on established syntactic properties. As put forth by Mahootian and Santorini (1996), the Complement/Adjunct Distinction is seen in (24):

(24) *Heads determine the syntactic properties of their complements in code-switching and monolingual contexts alike.* (=14 in Mahootian and Santorini 1996)

They emphasize that the effects of the CAD can be observed in monolingual language, unlike the

Functional Head Constraint. A switch that is permissible according to the CAD can be seen in the attested example in (25) while the invented example in (26) is not permitted.

(25) *Wahed* une cousine
One a cousin
'One cousin' (=71 in Bentahila and Davies 1983)

(26) **wahed* cousine
**One* cousin
'One cousin'

The code-switch seen in (25) is permitted by the CAD because the Arabic Determiner *wahed* selects for a full DP. For this same reason (26) would not be allowed by the CAD; here *wahed* is followed directly by a noun. It is only the syntactic properties of the employed lexicon that are determined; there is no language selected for in any way. Adjuncts are considered to be unrestricted by this account.

The CAD can easily be tested in the Determiner Phrase. A switch between a Determiner and its complement should be allowed, but is only permissible if the structure of the switched complement is what the Determiner would select in a monolingual DP. The structure of an adjunct is not selected for in any way; any adjunct would be allowed by this constraint. The CAD is therefore only visible in the code-switched DP if the languages in question take different structural complements to the DET. As will be seen in section 2.2.2, this is the case between French and the dialects of Arabic.

An especially interesting point made by Mahootian and Santorini comes from their evidence against the Functional Head Constraint. They give several examples of Moroccan Arabic-French switches in which a switch occurs between a Moroccan Arabic demonstrative determiner and a French definite determiner, seen in (27):

(27) a. *dak* la chemise

that the shirt
'that shirt'
b. *wahed* le liquide
one the liquid
'some liquid'
c. *wahed* une cousine
one a cousin
'one cousin' (=(69)-(71) in Bentahila and Davies 1983)

All of the examples in (27) switch languages between a Determiner, a type of functional head, and its complement. Clearly, this demonstrates that the Functional Head Constraint is not universal. However, the lack of this type of data in Belazi et al.'s study may indicate that the structure is not permissible in some dialects of Arabic in contact with French, including Tunisian Arabic-French code-switching. Recall that Belazi asserts that the switches of a noun after a definite determiner are due to borrowing. This seems less likely to be the case for Moroccan Arabic as the items 'liquid' and 'cousin' may be hard to argue as borrowings to fill lexical gaps. If these switches are possible in Moroccan Arabic, but not Tunisian, it is evidence that the Complement/Adjunct Distinction overgenerates possible switch locations in that it cannot account for the lack of Determiner-Noun switches in the Tunisian data, indicating that the CAD is also not universal. It is for this reason that it seems the two dialects, Moroccan Arabic and Tunisian Arabic, may be code-switched in different ways with French. This possibility has not yet been explored and is the aim of the current study.

Mahootian's (1993) Null Theory has been expanded by the authors and receives more attention than the CAD and, while the two are similar, they are distinct. The Null Theory states:

(28) The language of a head determines the phrase structure position of it complements in code-switching just as in monolingual contexts.
(Santorini & Mahootian 1995)

The Null Theory limits only the position of the complements, not the internal structure of them.

The fact that the Null Theory has been further developed while the Complement/Adjunct Distinction seems to have been abandoned may indicate that the authors found that it is in fact only the position, and not the internal syntactic structure, of the complement that is relevant. This is seen in a Farsi-English example in (29)

(29) you'll buy *xune-ye* *jaedid*
you'll buy house-part new
'you'll buy a new house' (=106 in Mahootian 1993, 152)

The English verb *buy* selects for a Determiner Phrase. In a monolingual English sentence, this would begin with an article, *a* in (29). However, the Farsi DP does not have the same structure. This switch is allowed because the code-switched complement is a valid DP in Farsi, despite the fact that it does not follow the syntactic structure of English. If it is only the position that is selected for, the stronger hypothesis of the CAD should be shown false in that the syntactic structure of a complement is not determined by the head.

2.1.4.3 The Functional Head Selection Constraint

In 2003, Chan proposed a modified version of the Functional Head Constraint that also incorporates some aspects of the Complement/Adjunct Distinction. He calls this the Functional Head Selection Constraint (FHSC), which states that:

(30) Code-switching can take place between a functional head and its complement provided that the c-selection restriction of the functional head is observed.
=(14) in Chan (2003, 151)

The FHSC diverges from the FHC by not positing assuming the existence of a language feature on all lexical items. Instead, similar to the CAD, he states that it is only the syntax of the switched complement clause that must obey that of the functional head. As long as the complement is of the syntactic or phrasal category selected by the funtional head, it is allowed as in (31)

(31) je peux le dire *had* le truc *hada* *baš* je commence à apprendre
I can it say this the thing here that I begin to learn
"I can say this in order that I start to learn."
(=101 in Bentahila and Davies 1983, 323)

Again there is no specification of language, only for the structure. The type of functional head found in the DP used in the present study is the Determiner. This is also the only element of the DP in French or Arabic that requires a complement. For this reason, the CAD and the FHSC give the same permissible and prohibited switch types within the Determiner Phrase.

Outside of the Determiner Phrase, the differences between these two constraints can be seen. The FHSC does not constrain switches between a lexical head and its complement, such as a verb or noun, allowing switching between lexical heads and their complements irrespective of the syntax of the switched clause.

2.1.4.4 The Matrix Language Frame Model

The Matrix Language Frame (MLF) Model was originally introduced by Myers-Scotton in 1993 and has since been refined many times with the most recent addition being the 4-M Model (Myers-Scotton & Jake 2009.) It differs from the models listed above as it does not rely on a particular syntactic theory. Instead, Myers-Scotton defines her own categories to explain the observed patterns in code-switching. The main proposal of the MLF is that the two languages used in code-switching are not equal. One is considered the Matrix Language (ML), and is syntactically the primary language of the utterance. The other language is the Embedded Language (EL). This idea is similar to that first put forth by Joshi (1985) who first suggested the idea of an ML and EL. He defined the ML as the language that the interlocuteurs identify as the "coming from" language (191). He also emphasized the asymmetry between the two languages

as well as constraints on where switches may occur. Importantly, his model prohibited switching of certain closed class items and complementizers. Myers-Scotton takes a different approach by detailing her own classes of morphemes and identifying whether they are involved in code-switching.

It is important to recognize in considering the MLF model that Myers-Scotton has stated that it refers only to "Classic CS" (code-switching), which she identifies as "CS in which empirical evidence shows that abstract grammatical structure within a clause comes from only one of the participating languages" (2009, 337.) This seems to border on circular logic: the MLF, which depends on the notion that one language is the source of phrase structure, is valid only for Classic CS, or clauses that demonstrate that one language is the source of phrase structure. This makes it easy to define any example of CS that does not fit the expected pattern according to the MLF as non-Classic CS and therefore not subject to the principles of the MLF model.

Definitions of classic CS aside, three main principles underlie the whole of the MLF model. The Uniform Structure Principle serves as a basis and is expanded on by the Morpheme Order Principle and the System Morpheme Principle. The Uniform Structure Principle (USP) states that:

> (32) A given constituent type in any language has a uniform abstract structure and the requirements of well-formedness for this constituent type must be observed whenever the constituent appears. In bilingual speech, the structures of the Matrix Language (ML) are always preferred. Embedded Language (EL) islands (phrases from other varieties participating in the clause) are allowed if they meet EL well-formedness conditions, as well as those ML conditions applying to the clause as a whole (e.g. phrase placement.) (Myers-Scotton & Jake 2009, 337)

In this way, it is not a specific grammatical framework that is referred to, but a more general notion of "well-formedness." The Morpheme Order Principle adds:

> (33) Morpheme order must not violate ML morpheme order (Myers-Scotton 1993, 7)

The Morpheme Order Principle indicates that the morpheme order is from the ML, but the model also allows for the possibility of so-called "EL Islands," defined as "EL phrase-level constituents" (Myers-Scotton & Jake 2009, 337). It is the final principle, the System Morpheme Principle, that relates directly to the current study. Myers-Scotton describes system and content morphemes in terms of quantification and ability to assign or receive thematic roles (Myers-Scotton 1993). Any lexical item that can quantify others is a system morpheme, such as quantifiers, tense/aspect and determiners. The second type of system morphemes are not quantifiers, but cannot either assign or receive theta-roles, such as complementizers, copulas and dummy pronominals. It is for this reason that the current study cannot evaluate the MLF in any way. The System Morpheme Principle states:

(34) One type of system morpheme must come from only one of the participating language and this language is identified as the ML. (Myers-Scotton 2010)

The 4-M model is the most recent addition to the MLF, detailed in her 2009 article. The 4-M model identifies four types of morphemes: Content morphemes, early system morphemes (SMs), Bridge late SMs and Outsider late SMs. The MLF model is proposed to work only for Outsider late SMs. Nouns and adjectives are both content morphemes, placing them outside the realm of the MLF model. Determiners, which had in the past been a source of criticism of the MLF, are specifically stated by Myers-Scotton to be a type of early SM (Myers-Scotton 2010). For this reason, they also are considered to not be constrained in the MLF model. As none of the potential elements of a French/Arabic determiner phrase is an Outsider late system morpheme, the model cannot be applied to the DP and for this reason does not apply to this study. This overview of the MLF is included due to its recent influence in the field of code-switching. It will not be considered further as the syntactic structures evaluated in this study are outside of its

scope.

2.1.4.5 Muysken's Bilingual Speech

As the theories described above, like all others postulated for code-switching, have been the object of major criticism, Muysken (2000) proposes that all may be incomplete, but at the same time complement each other. As code-switching is doubtlessly affected by societal factors in addition to the syntax of the languages involved (citations), he theorizes that the types of code-switching found between language pairs are affected by the social situations in which they are employed. For this reason, it is likely that switching between the same language pair in different social contexts may affect the syntactic locations in which code-switching occurs.

The three categories of code-switching identified by Muysken are Insertion, Alternation and Congruent Lexicalization. Each of these categories is similar to one or more theories of code-switching put forth in the past but has since been demonstrated to be inadequate for data from another language or group of languages. A specific speech community may prefer one of the types of switching over others, to the extent that only one type is found, or a combination of the three may exist within a single speaker or group of speakers.

Insertion is the pattern in which one main language supplies the structure into which a word or entire constituent from the other language may be included. Muysken views the MLF as a model that deals mainly, if not solely, with insertion. Some grammatical constraints would also be types of insertion. This pattern is summarized as ABA, demonstrated in (35).

(35) Yo anduve *in a state of shock* por dos días.
I walked *in a state of shock* for two days. (Muysken 2000, 5)

Depending on the language pair, an inserted unit may be as small as a word or a constituent

such as a noun phrase or adverbial phrase. Muysken considers borrowing a type of Insertion limited to a single lexical item. He does not limit Insertion to a specific direction in code-switching; it may happen between any language pair in either directions. Muysken states that Insertion is commonly found in situations of asymmetric bilingualism, such as colonized areas or the language of recent migrants. This situation accurately describes that of Tunisia and Morocco. The countries are no longer governed by France, but the use of French stems from the colonial period.

Alternation can be seen in language pairs such as Spanish-English, where there is relative compatibility of the two grammars, at least at switching locations. The two languages involved occur alternately, with each language retaining its own structure. It can be considered that "one language is replaced by the other." (Muysken 2000: 5)

(36) Andale pues *and do come again.*
That's alright then *and do come again.*
(Gumperz and Hernández-Chavez 1971: 118)

Alternation is similar to Poplack's Equivalence Constraint and, according to Muysken, is likely to occur in stable bilingual communities where languages are traditionally separated. In such a community, code-switching may be less common or have a more negative association. This is the pattern found in the stimuli of the experimental section of this study, and in some ways the social setting of Tunisia and Morocco may match that proposed by Muysken. It is true that French and Arabic are more closely associated to different domains, such as French with technology and modernity and Arabic with religion and tradition (Bentahila 1983a, Benrabah 2007, Lawson and Sachdev 2000) but they are also often used in many of the same domains (Lawson and Sachdev 2000). If the languages are commonly mixed, not strictly separated, the pattern of Alternation may be less likely for bilingual Arabic/French speakers.

Congruent lexicalization is somewhat more complex than the other two types of code-switching. Congruent lexicalization occurs when the grammatical structure of the languages involved is shared, but the lexical items come from more than one langauge. This could be due to grammatical convergence or similarities between languages. The monolingual varieties being used must be well understood by the speaker in order to switch with ease between them. There are several contexts in which congruent lexicalization may be found. Muysken states that this may be more common in style-shifting between different varieties or registers of a single language. Another is between closely related languages. An example between Dutch in English is in (37)

(37) Weetjij [whaar] Jenny is?
'Do you know where Jenny is?' (=7 in Muysken 2000:5)

As Muysken indicates, the structure of the sentence is similar in both languages, and the phonology is as well. *Whaar*, when pronounced by a bilingual, may be hard to distinguish as either Dutch or English, with the pronunciation of the name and *is* being homophonous in the two. A third situation is one in which the languages have the same level of prestige or where the languages are not traditionally separated. While he generally associates Poplack's Spanish-English code-switching examples with alternation, sometimes he finds congruent lexicalization a more satisfying description, as in (38).

(38) (A) Why make Carol *sentarse atras* (B) *pa'que* everybody has to move
sit at the back so that
(C) *pa'que se salga.*
so that she may get out (Poplack 1980: 589)

While he notes that this could be seen as an example of multiple insertions, the fact that switching is found within a constituent with no apparent relationship between the fragments in each language makes this a less appealing proposal. If the target community of bilingual French

and Arabic speakers does not traditionally separate languages, it is possible that congruent lexicalization may also be found in the dialects of the present study. In Bentahila and Davies's 1995 study, this may have been the case for the older generation, who grew up with French and Arabic both used often in daily life and who Bentahila and Davies describe as balanced bilinguals. The older speakers were found to switch back and forth from Arabic to French at a wide variety of syntactic locations. However, this situation can change quickly as the younger generation was found to switch mainly for a whole NP. This generation of Moroccans was educated after many Arabization policies were put into place, giving French a more marginal role in their education and lives.

2.2 Background on target languages and societies in which they are spoken

2.2.1 Varieties of Arabic examined in the current study

This study will examine code-switching between French and two different varieties of Arabic: Moroccan Arabic and Tunisian Arabic, hereafter also referred to as Moroccan and Tunisian. These particular dialects have a history of contact with French, dating back to the 1880s or longer, and French continues to be widely used in Morocco and Tunisia for academic, personal and professional reasons (Ennaji 1991, Bassiouney 2009, Bentahila and Davies 1983a, b, Lawson and Sachdev 2000, etc). Due, in part, to sustained contact, French-Arabic code-switching is a common phenomenon. The Moroccan and Tunisian dialects of Arabic are classified as Magrebin dialects (along with Algerian), as defined by Versteegh (1997). However, the mutual intelligibility of the dialects that form this group is questionable. Ennaji states that they are "usually mutually comprehensible in the sense that a Moroccan Arabic speaker can generally communicate with a Tunisian or Algerian Arabic speaker without any loss of meaning."

(1991: 11) His careful use of ‘usually’ and ‘generally’ implies that the situation is not straight-forward. Referring to speakers of all Arabic dialects, Mahmoud (1986) bases mutual intelligibility on the conditions of geographic distance and the level of education of the interlocutors. He states that smaller geographic distance and higher education levels lead to higher comprehensibility. As dialects often vary by location, the stipulation of distance is unsurprising. The factor of education, in contrast, would not be expected to matter for intelligibility of Arabic dialects. The dialects are not taught in schools and in many areas are not used in schools at all. For this reason, if more educated individuals understand each other more easily, it is likely due to use of an educated variety of Arabic that is separate from the individual dialects in question, not due to mutual intelligibility of the dialects.

A classic description of language use in Arabic-speaking countries is diglossia. Ferguson coined the term diglossia, based on the French term *diglossie*, in his seminal article in 1959. He defined diglossia as,

"a relatively stable language situation in which, in addition to the primary dialects of the language (which may include a standard or regional standards), there is a very divergent, highly codified (often grammatically more complex) superposed variety, the vehicle of a large and respected body of written literature, either of an earlier period or in another speech community, which is learned largely by formal education and is used for most written and formal spoken purposes but is not used by any sector of the community for ordinary conversation." (1959, 336)

Arabic is one of the four languages that Ferguson used to illustrate diglossia but, as Walters (2003) points out, the situation described by Ferguson may have been more idealistic than reflective of use in any Arabic-speaking country from the beginning. In Morocco and Tunisia, as in other Arabic-speaking countries, Modern Standard Arabic (MSA), or Fusha, certainly is highly codified with a respected body of written literature. The use of MSA, Classical Arabic and

Dialectal varieties leads Ennaji (1981) to call the Maghreb triglossic, though this does not take into account the use of non-Arabic varieties, including Berber and European languages.

Walters (2003) describes Tunisia as a post-diglossic society due to the many different varieties of language used. MSA continues to be used in religious contexts, in official written documents, and in newspapers. Still, use of the dialects is becoming more common in areas that were once the domain of MSA (Mahmoud 1986), including in highly formal contexts as exemplified in speeches by former present Bourguiba (Walters 2003) and current president Ben Ali (Bassiouney 2009). In formal speaking situations, MSA is used in order to demonstrate the speaker's command of the language, but the dialect may also be used to increase comprehensibility for the target audience. For conversations in formal settings, MSA may be used at the beginning, though the conversation often incorporates an increasing amount of dialect as the exchanges continue. Other languages or varieties of Arabic may be used based on the dialects spoken by other interlocutors and the speaker's comfort level with Standard Arabic as well as European languages. Instead of a diglossic situation, with two varieties used in distinct contexts, many varieties are used in the same social contexts.

2.2.2 The Determiner Phrase in Arabic and French

It is critical to understand the structure of the determiner phrase (DP) in both Arabic and French separately before considering how they might be combined. The DP in the dialects of Arabic being studied has the same basic form, but minor differences exist between them. The similarities and divergences between the Arabic and French DPs make it possible to test the proposed code-switching constraints described above.

2.2.2.1 The Determiner Phrase in Arabic

The basic form of the DP in the dialects of Arabic being studied is Determiner Noun (Adjective). The Adjective is optional, noted using the standard syntactic notation of parentheses. This is conventionally written as DET N (A). This structure can be seen in (39).

(39) *l-karhba mizyaena*
the-car nice
"The beautiful car" (Tunisian, from Belazi et al. 1994)

In MSA, when the noun begins with a coronal consonant, the definite article 'l' assimilates to the initial consonant of the noun (Holes 2004), as in (40)

(40) r-rajul
'the man'

MSA also allows bare nouns, which are indefinite, as in (41)

(41) rajul
'a man'

Most Arabic varieties require that an article be used with an adjective modifying a definite noun, seen in (42), but below it will be seen that this does not apply in the two dialects examined here.

(42) it-taye:r il-mo:v
the-outfit the-mauve
'the mauve outfit' (Egyptian Arabic, from Brustad 2000, 41)

One possible difference between Arabic and French, case, does not need to be considered. While Classical and Modern Standard Arabic use nominal case markings, the nominative and accusative cases are syncretic in the colloquial dialects.

2.2.2.1.1 The DP in Moroccan Arabic

Lahlou (1991) gives a description of the DP in Moroccan Arabic. The basic DP follows the

same form indicated as in 1.2.1, with the possible structure of DET N (A), and assimilation of the definite article to initial coronal consonants of the following noun. The standard form can be seen in (43) with assimilation in (44).

(43) l-bent
the.girl

(44) d-dar
the.house

When the definite article assimilates, as Lahlou explains, a geminate consonant begins the word. This is the same as in MSA.

While not possible in MSA, another form of the Determiner Phrase that exists in Moroccan Arabic is: Demonstrative Determiner Definite Determiner Noun (Adjective). As the two types of Determiners have very different syntactic behaviors and requirements, the Demonstrative Determiner will be abbreviated at DEM and the Definite Determiner will be abbreviated as DEF. The abbreviation DET will be used only when discussing all determiners. The structure of DEM DEF N can be seen in (45) without an adjective.

(45) *Dak l fqi […]*
that the teacher […]
"That teacher […]"

This structure only occurs in Moroccan Arabic when the first determiner is the word 'wahed' meaning 'one' (indefinite), or when the first is a demonstrative determiner. 'Wahed' is seen in (46) and a demonstrative determiner, *had* in (47).

(46) wahed l-bent
one the-girl
"one girl"

(47) had l-bent
this the-girl
"this girl"

The DP can also be represented in a tree diagram, sentence (47) is seen in Figure 1 following the structure given by Bentahila and Davies (1983) with the head considered to be the Determiner, as assumed by Belazi (1991) based on evidence from Abney (1987).

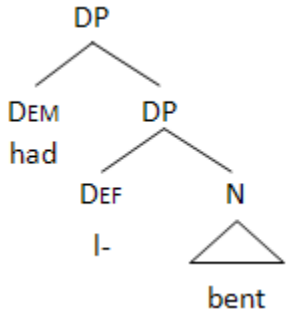


Figure 2 – Moroccan Arabic Determiner Phrase with two determiners

A second indefinite marker is sometimes identified in Moroccan Arabic, 'ʃi.' Bentahila and Davies consider these to have different meanings, translating 'wahed' as 'one' and 'ʃi' as 'some.' Brustad confirms that they are distinct, though considers neither a true indefinite. Instead, she indicates that both are markers of different levels on a spectrum of definite/indefiniteness.

According to Brustad (2000), there are three possible forms of demonstrative determiners in Moroccan Arabic -- proximal, distal and unstressed. These are seen in Table 1 below.

Table 1: Demonstrative Pronouns in Moroccan Arabic

	masculine	feminine	plural
Proximal	hāda	hādi	hādu
Distal	(hā)dāk	(hā)dīk	(hā)dūk
Unstressed	hād	hād	hād

All types of demonstratives require the co-occurrence of a definite article, but it is the unstressed demonstrative form, 'had,' that is seen in sentence (39) and considered in the present study. As can be seen in Table 1, /hād/ has only one form and does not agree with the number or gender of the noun it modifies. In describing it further, Brustad sites Harrell (1962), who explained that the unstressed /hād/ has "none of the implications distinguishing between

near and far as do the English demonstratives *this* and *that*" (p. 147). In this respect, /hād/ is similar to the French *cette*, which cannot distinguish distance. Brustad identifies /hād/ as an anaphoric demonstrative. This is due to the fact that they are used only "to mention entities whose unique identity is already known to both speaker and listener" (p.148). The anaphoric demonstratives can be used in similar ways to definite articles. However, there is a pragmatic distinction between the two. Anaphoric demonstratives give greater emphasis to the nouns they modify, indicating a more important role in the discourse.

2.2.2.1.2 The DP in Tunisian Arabic

The basic structure of a DP in Tunisian Arabic is the same as that given in section 1.2.1, DEF N (A). Belfalah (1992) gives a more detailed description of the DP, which he classifies as an NP, in Tunisian Arabic. Like in MSA, a bare noun may stand on its own to constitute a DP, or a pronoun may be the only element present. These two possibilities are seen in (48) and (49).

(48) -bàb
a door

(49) hiyya
she (from Belfalah (1992) p.68)

As noted above, a determiner may also precede a noun, as seen with a definite determiner in (50). Like in Moroccan, some determiners require the co-occurrence of a definite determiner. In Tunisian Arabic, demonstrative determiners, quantifiers and numerals all fall into this category, as seen with a demonstrative in (51).

(50) DEF N
lqasm
the.class

(51) DEM DEF N
hàk lqàsm
that the.class ('that class')
(from Belfalah p.68)

Belfalah distinguishes two classes within demonstratives. Class A demonstratives show gender

and number agreement with the nouns they modify. They can also create what Belfalah calls a noun verbal sentence when followed by an indefinite noun, elsewhere called an equational sentence, as no copular 'to be' exists in Arabic, as in (52) and can also be used after a definite noun, as in (53).

(52) hàḍa ktàb
this/that book
"this/that is a book"

(53) lktàb hàða
the book this/that
"this/that book"

Note that in (44) above the demonstrative is used directly preceding the noun, with no intervening definite determiner. It is only in contexts such as (45), where the demonstrative follows the noun, that a definite determiner can modify the same noun as Class A demonstratives.

Class B demonstratives have a similar function to Class A with a few important differences. Unlike Class A demonstratives, a proximity distinction is also made that corresponds to "this" and "that" in English, with the form 'ha' meaning 'this' and 'hak,' meaning 'that'. The form used in this study is 'ha' as it was considered more appropriate by the native speaker. Class B demonstratives must be used with definite nouns, a definite determiner followed by the noun, as seen in (54) below.

(54) (a) ha lktab
this the.book
"this book"

(b) ha lktub
these the.books
"these books"

As visible in (54a and b) 'ha' does not agree in gender or number with the noun it modifies, similar to the Moroccan form 'had.' No information is given to indicate whether Tunisian Class B demonstratives are anaphoric in the same way as the Moroccan 'had.'

While the two use different lexical items, the basic syntactic structure of the DP in

Tunisian is the same as that in Moroccan. There may be a subtle difference between the demonstrative determiners in the two dialects as it is not certain whether the Tunisian 'ha' has an anaphoric function. Aside from this point of potential difference, there is no syntactic reason why the two should be used in code-switching with French in different ways. This similarity means that the different permitted and prohibited switch types according to the FHC and CAD, explained above, are not due to syntax, but another factor. If syntactic constraints are universal and have as much influence on switch types as proposed by these authors, other factors should not strongly affect the acceptability of switch types. It is then the lack of syntactic difference between the two dialects and the contradictory nature of the constraints on CS based on data from the dialects that underlies the importance of exploring code-switching between each dialect and French separately.

2.2.2.2 The DP in French

The structure of the Arabic DP is similar in many ways to that of French. The basic form of the DP in French is DET N (A) and can be seen in (55).

(55) La boîte rouge
The box red
"The red box"

Most adjectives in French are post-nominal, but there is a small class that is pre-nominal. When one of these adjectives is used, the general form becomes DET (A) N, as in (56).

(56) La jolie boîte
The pretty box
'The pretty box'

The situation becomes further complicated by the fact that, while the adjective "jolie" and all

other pre-nominal adjectives are generally placed before the noun they modify, Bouchard (1998) states that they can also be used post-nominally. He suggests that all adjectives may be used in either position, with a change in semantics when the adjective is found out of its canonical position. Thus, any adjective can follow a noun in French. In French, demonstrative determiners are followed directly by the noun they modify and are never followed by an article, as seen in (57).

(57) a. Cette boîte rouge
This box red
"This red box"

b. *Cette la boîte rouge
*This the box red
*"This the red box"

2.2.2.3 The combined DP in French-Arabic code-switching

The differences between the two languages, then, are that French requires a DEM to be followed by an N, whereas the dialects of Arabic require a DEF to follow a DEM. The dialects of Arabic also do not allow prenominal adjectives. Having established the structure of the determiner phrases, the theoretically possible French-Arabic and Arabic-French DPs can be identified. The Arabic/French determiner phrase yields seven possible sites where switches may theoretically occur, making 14 possible types of switches, seven beginning in Arabic and switching to French, illustrated in (58), and seven that do the opposite, as in (59). These sites are: between a demonstrative determiner and definite determiner (58a, 59a), between a demonstrative determiner and noun (58b, 59b), between a definite determiner and noun (58c, 59c), between a noun and postnominal adjective (58d, 59d), between a prenominal adjective and noun (58e, 59e), between a definite determiner and prenominal adjective (58f, 59f), and between a demonstrative determiner and a prenominal adjective (58g, 59g). All examples in (58)

and (59) are in Moroccan Arabic-French.

Code-switches from Arabic to French

58 a) *had* la belle ville
this the beautiful city
'This beautiful city'

b) *had* ville
this city
'This city'

c) *l* ville
The city

d) *l medina* sale
the city dirty
'The dirty city'

e) *l zwina* ville
the beautiful city

f) *l* belle ville
the beautiful city
'The beautiful city'

g) *had* belle ville
this beautiful city
'This beautiful city'

Code-switches from French to Arabic

59 a) Cette *l medina zwina*
this *the city beautiful*
'This beautiful city'

b) Cette *medina*
this *city*
'This city'

c) La *medina*
The *city*

d) La ville *mouskha*
The city *dirty*
'The dirty city'

e) La belle *medina*
The beautiful *city*

f) La *medina zwina*
The *city* *beautiful*
'The beautiful city'

g) Cette *medina zwina*
this *city beautiful*
'This beautiful city'

The possible syntactic permissibility of the DPs varies greatly between these locations. 58a and 59a correspond to the required use of determiners in the dialects of Arabic, while 58b and 59b reflect the necessary elements and order in French. 58c and 59c have the correct word order for both languages, but the switch occurs between a functional head and its complement, an oft-debated location for a switch. 58d and 59d also follow standard word order in both languages, with a switch occurring between a lexical head and its adjunct. 58e, f, g and 59e, f, g display the proper word order for French when using the adjective "zwina/belle" 'beautiful,' a pre-nominal adjective. As the dialects of Arabic being examined do not allow pre-nominal adjectives, the placement of this adjective in an Arabic phrase should be after the noun. The possible variation in adjective placement in French, even if rare, could mean that "belle" would

be accepted post-nominally in a switched sentence. In order to avoid the potential ambiguity of pre-nominal French adjectives only post-nominal adjectives will be considered in the current study.

This leaves four possible switch locations, seen in (58a-d) and (59a-d) above. The invariable form of the Arabic demonstratives used in the study means that gender agreement between the demonstrative across languages is not an issue when a sentence begins in Arabic. Similarly, the definite article in Arabic has a single form. For this reason, the only possible circumstance where a switch may potentially disagree in gender or number is between a French DET and an Arabic N, as in (59a, b, c) and between a noun and adjective, in (58d) and (59d). Therefore, although a limited set, the four target switch locations identified make it possible to test the FHC, CAD and FHSC with a relatively restricted number of possible complications.

2.2.2.3.1 Implications of the Functional Head Constraint on the French-Arabic DP

When applied to the possible French-Arabic DP, the Functional Head Constraint disallows switching between a DET and its complement. This should apply to both demonstrative and definite determiners, as both are heads. Specifically, the switches seen in (58a-c) and (59a-c) should not be allowed while those in (58d) and (59d) should be permitted. The Word-Grammar Integrity Corollary, proposed with the FHC, regards switches here between a noun and an adjective that modifies it. More specifically, in the possible DPs examined, it means that switches between a noun in one language and a postnominal adjective in the other, in (58d) and (59d) is allowed because both languages contain this structure. By creating stimuli with and without switches in these syntactic locations, the acceptability of both the Word Grammar Integrity Corollary and the Functional Head Constraint can be tested within the determiner

phrase.

2.2.2.3.2 Implications of the Complement Adjunct Distinction (CAD) on the French-Arabic DP

The only type of heads explored in the current study that take a complement are DEM and DEF. As definite determiners select the same type of complement in both languages, a noun, the CAD would state that a switch between the two, (58c) and (59c), is acceptable. As an Arabic demonstrative selects a definite determiner and a noun, this should remain true when a switch follows the demonstrative if the CAD is correct. In the same vein, the French demonstrative determiner should be followed only by a noun. Thus, the examples in (58a) and (59b) would be grammatical according to the CAD, while those in (58b) and (59a) would not. This model does not constrain switches between a noun and a post-nominal adjective, meaning that (58d) and (59d) should be allowed.

2.2.2.3.3 Implications of the Functional Head Selection Constraint on the French-Arabic DP

The functional heads explored in the current study are the demonstrative and definite determiners. The distinction made by the FHSC for the current data, then, is the same as that made by the CAD. Specifically, (58a, c, d) and (59b, c, d) should be acceptable, while (58b) and (59a) should not. While these two constraints have different effects outside of the DP, within it none can be observed.

2.3 Social variables and code-switching

Like many other linguistic phenomena, there are many factors including sex, socioeconomic status, and region that play a role in the acceptability of code-switching as well

as the amount of code-switching a person produces. While all of these are valid and must be taken into account in order to understand the variety found in spoken language, only the variable of sex will be considered in the present study. In most cultures, females tend to speak more standard forms, particularly those with overt prestige, while males use more non-standard forms that tend to carry covert prestige (J. Milroy 1981, Labov 1991, Wodak and Benke 1997). Ibrahim (1986) observes that many studies have indicated that Arabic women tend to use more non-standard forms than men, the opposite of most language groups. He also warns that identifying the standard and prestige varieties for Arabic is often based on Ferguson's (1959) notion of diglossia. Diglossia, as discussed in section 2.2.1, presumes that the High variety of Modern Standard Arabic (MSA) is also the sole prestige variety in a community. Walters' (2003) research in Tunisia indicates that Diglossia may not be an accurate description of modern Arabic-speaking societies in which multiple varieties of Arabic, in addition to European languages, are spoken. If women are striving to use forms with overt prestige, it is possible that this prestige may be found in European languages, here French, in addition to in MSA. As Lawson and Sachdev (2000) note, in Tunisia it is perceived that women switch more, while a matched-guise study revealed that men are seen as having more prestige when code-switching between French and Arabic than women who do so.

As the stimuli used in the experimental portion of this study were recorded by male speakers, the ratings of the switch types may be higher than it would have been if females were recorded. If female speakers indeed participate in code-switching more often than males, they could be more sensitive to any syntactic constraints on CS that exist. For these reasons, sex may greatly affect how a speaker participates in and approves of code-switching and must also be examined.

2.3.1 Social variables: A brief description of the linguistic situation in Morocco and Tunisia

Code-switching is undoubtedly a social phenomenon. For this reason, it cannot be assumed to operate in a vacuum, irrespective of social norms. It may be that there are some syntactic constraints that are universal or apply to typologically similar language pairs, but this should require evidence and therefore not be simply assumed from the outset. Consequently, the social contexts of the countries examined here is necessary. The countries in which the target dialects are spoken were once under the control of France. However, the extent and duration of this control differs, leading to diverging use of French today. The attitudes of the populations and actions of the governments on gaining independence from France also vary greatly. This has meant that the present societies and education systems use French in different ways, perhaps causing even more divergence in the prestige of the language in the countries involved in this study. A language with higher prestige will likely be learned by a larger proportion of the population than a language with lower prestige, leading to a large portion of the population with a higher level of proficiency in a prestige variety.

Social variables affect the likelihood that a speaker will participate in code-switching, the direction in which she code-switches and her exposure to others who code-switch. Naturalistic data would be most reflective of these factors, while elicited judgments abstract away from social variables somewhat by proposing possible sentences instead of analyzing attested utterances spoken by a specific individual in a specific context. Nevertheless, judgments from a speaker are inherently based on her experience with code-switching in the past, which is itself predicated on social factors. it is for this reason that social factors are discussed here.

2.3.1.1 Linguistic situation in Morocco

The linguistic landscape of Morocco is likely the most complex in North Africa. As in other parts of the region, Arabic is the official language, while Berber is also spoken by an estimated 40% of the population (Ennaji 2002) French is also widely used. The most recent census figures from the Moroccan government, 2004, indicate that 89.8% of the population over 5 years of age reports speaking Moroccan Arabic, with 69.1% report being able to both read and write French (Morocco 2004). No figures are given for those who speak, but do not read and write, French. While it lacks any official status, French continues to be the 'elite language' (Bassiouney 2009). Spanish is also spoken by "a considerable population" (Scipione and Sayahi 2005) in the north of the country due to the proximity of Spain and frequent contact between the populations over centuries.

France's control of Morocco was short compared to other countries in the Maghreb. A large portion of Morocco was a French protectorate for 44 years, beginning in 1912 and ending in 1956. During these years, Morocco was divided into Spanish Morocco, in the north, and French Morocco, in the south. The protectorate lasted long enough for the colonizers to impose French as the language of education, government and the media, but not to create a bilingual Moroccan population. One reason for this is that the population centers for Moroccans and colonizers were different. The small population of colonizers lived on the coast, while the Moroccans generally lived inland. This led to few Moroccans having a need to speak French while the country was a protectorate. A government census in 1960, the first year figures are available, indicates that just over 7% reported speaking French and 6% reported being literate in it. (Sirles 1999, Morocco 1960). This changed after independence due to inconsistent language policy (Sirles 1999). One area that had a large impact on the spread of French is education.

Increased access to education was highly valued after independence, with an emphasis on Moroccan instructors. However, most of these instructors were trained in France and used French to teach their subjects. In this way, use of the language among Moroccans actually increased after independence with French used throughout the country despite efforts for Arabization.

In the current educational system, Arabic is used as the language of all content courses at all levels with French as the first second language. Moroccan students begin learning French in the third year of schooling, and take 8 hours of the language per week. This gradually is reduced over time to 4-5 hours in the final two years of schooling. In addition to being the language of content courses, students also study Arabic for 6.5 hours per week early on, decreasing to 1-5 hours per week during the final two years, depending on the student's chosen track. A second foreign language is only offered in the 10^{th} to 12^{th} years, indicating the privileged status that French maintains. Moroccan students do not use French in primary or secondary school to study other subjects such as math or science; French is only a foreign language and not a vehicle for teaching. However, at the university level, French remains the language of instruction for medicine, engineering and sciences. (Bassiouney 2009.) While all individuals who have completed secondary schooling have had a large amount of exposure to French, students who study subjects taught in French at the university level may have greater proficiency in French than those who do not. Proficiency has been shown to play a role in the use of code-switching between Arabic and French (Bentahila & Davies 1995). For this reason, this study aimed to recruit native speaker respondents of the dialects that consider themselves fluent in French.

The socioeconomic situation in Morocco also encourages the continued use of French.

Morocco's economy relies in part on tourism, making French and English important skills. Morocco's largest trade partner is France, creating a need for French among business people. It is hard to determine whether the language ability of Moroccans led to this status or vice versa. On independence, the elites sent their children to French schools. This may have been the beginning of French as a requirement for high-level positions, entrenching the status of French as an elite language. As demonstrated by Ennaji (2002), university students confirm that French continues to have a favored status, as 73% of those he surveyed were in favor of French-Standard Arabic bilingualism, as were 78% of the professors surveyed. Their opinions are not surprising, given that most of the political and business leaders in France receive some or all of their university education in France (Sirles 1999).

2.3.1.2 Linguistic situation in Tunisia

The linguistic situation in Tunisia has some similarities with that of Morocco as the two countries gained independence within a month of each other. Tunisia was under French control for 75 years, close to twice as long as Morocco. Colonized by France in 1881, it was occupied until 1956. Gallagher (1964 – thru Sirles) estimates that about 40% of the Tunisian population spoke French on gaining independence. The languages spoken before the arrival of the French and the policies after independence have made the current status of French markedly different. In Tunisia, Arabic and French are the two main languages, with English slowly becoming more widespread (Walters 2003, Bassiouney 2009). The Berber population makes up only about 1% of the country with Spanish and Italian used by small portions of the population. The president of Tunisia stressed the desirability of maintaining ties with France even as his country established its independence in 1956. Strong connections with France have continued to the present.

Arabization was considered important after independence, but the policy was not consistent. In spite of changing governmental whims, the country is now basically Arabized, but to a lesser extent than the others in North Africa. In 1999 the government set a goal to Arabize the entire administration by 2000. Today the government uses Arabic, but many necessary forms for daily activities, such as medical and banking documents, are still in French. Many large stores issue receipts only in French, emphasizing its status relative to finance. (Bassiouney 2009)

The educational system of Tunisia is more Francophone than other Arabic-speaking countries. Arabic is used from the first year with French being introduced as a foreign language in the third year. It is heavily emphasized with 9-11 hours per week during the first four years that students study it, though this drops to just 2-4 hours per week in the 12th and final year of schooling. The hours of Arabic instruction during the years of French instruction is 4.5 and 0-5, respectively. English is introduced here earlier than in Morocco, during the 7th year. Tunisian students have an even greater exposure to French within the school setting than these figures show; in secondary school Mathematics and all types of sciences are taught in French, while only History and Islamic studies, and Arabic are taught in Arabic. This is logical in a sense as math and sciences are the subjects taught in French at the university level. It also means that Tunisian students have a broader type of contact with French.

2.3.2 Implications of social variables on French-Arabic code-switching

The impact of social variables on code-switching between French and the dialects of Arabic is not as easy to identify as the predictions made by the code-switching models examined. It is clear that the populations of Tunisia and Morocco have different exposure to French throughout their education and for this reason may have differing levels of proficiency in

French. However, there are no exact predictions that can be made based on this fact. This study focuses only on the general question of whether speakers of the dialects prefer code-switching in general and by type of code-switch to differing degrees. The precise effects of social factors, while a topic that merits attention, is beyond the scope of the present study.

3. The study

3.1 Research questions and hypotheses

The aim of this study is to identify the switch locations that speakers of each dialect prefer and then to compare these locations to the models put forward in order to evaluate whether the models accurately predict these preferences. As numerous counter-examples have been found for all of the models described in section 2 except for Muysken's Bilingual Speech Model, it is not predicted that any will be shown valid for all language pairs in question. Instead, the goal is to investigate whether the same switch locations are preferred across dialects. Once this is determined, it will be possible to match these observed preferences with the models that predict them as well as with naturalistic data from each language pair. The research questions specifically are:

1 - Do speakers of each dialect show the same level of preference for code-switching overall?

2 - Do speakers of each dialect show the same preferences for type of switch?

3 - Do speakers of each sex show the same preferences for type of switch?

4 - Do the code-switching models accurately predict speaker preferences?

If the dialects share preferences, visible in the authenticity ratings that they give each switch type, this could indicate similar, if not identical, grammars between the dialects. If the dialects do not share preferences, it could be an indication of different grammars. However, another possibility is that social factors play a large role, as proposed by Muysken's model. If social factors are highly influential on CS, the same syntactic structures may be switched differently in the two communities.

It is hypothesized that speakers of the dialects will show the same basic preferences for

switch type, but that the strength of authenticity ratings will vary due to the fact that the code-switching models developed in part based on data from the dialects make different predictions. The variable of sex in regard to code-switching continues to be unclear in Arabic sociolinguistics and is included for further exploration. It is expected that there will be differences between the sexes, although the direction and extent of the divergence cannot be guessed. Finally, as the two main models examined predict different outcomes, one will be shown to be more descriptively adequate.

The overarching goal of the current study is to make a comparison of CS between similar dialects in contact with the same language. As the dialects have the same grammatical structure in the DP, any differences in preferred switch type will be based on distinct uses of code-switching between each dialect and French. Of the five code-switching models discussed above, three make clear judgments regarding the possible code-switching locations considered here. These are seen in Table 2 below. Muysken's model indicates that preferred switch types reflect different linguistic and sociolinguistic situations and for this reason makes no definite predictions regarding the present data. As noted, Myers-Scotton's MLF model does not apply to any of the parts of speech used in the current study. Both Muysken's Bilingual Speech model

Table 2: Permitted and prohibited code-switch types by CS model

	Arabic-French				French-Arabic			
Model	Type 1	Type 3	Type 5	Type 7	Type 2	Type 4	Type 6	Type 8
FHC	N	N	Y	N	N	N	Y	N
CAD	Y	N	Y	Y	N	Y	Y	Y
FHSC	Y	N	Y	Y	N	Y	Y	Y

Each switch type begins in the first language indicated and changes to the second at a particular syntactic location. Switch types 1 and 2 occur between a DEM and DEF, 3 and 4 between a DEM and N, 5 and 6 between a N and A, 7 and 8 between a DEF and N. See section 2.2.2.3 for examples of each type.

and Myers-Scotton's MLF are not included in Table 2 as they cannot be evaluated in the same way. In this table, it can clearly be seen that the predictions made regarding the syntactic constraints on these language pairs are the same in several contexts. Specifically, the three models predict that switching between an Arabic Demonstrative Determiner and a French noun will be prohibited as the Demonstrative Determiners used in this study require a Definite Determiner to follow them. Similarly, they predict that a French demonstrative determiner followed by an Arabic Definite Determiner will be prohibited for the same reason; the French demonstrative selects for a noun. They all predict that switches between a noun and adjective to and from either language will be permitted as nouns are lexical, not functional, heads and adjectives are adjuncts, not complements. It can also be observed that the predictions made by the CAD and FHSC when restricted to the Arabic-French DP are the same; this is due to the fact that the functional head, the determiner, is also the only element in this phrase that requires a complement. Due to this fact, it is impossible to distinguish between the FHSC and CAD in the present context.

The table also shows switch locations where the FHC differs from the CAD and FHSC: between an Arabic DEM and a French DEF, between an Arabic DEF and French Noun, Between a French DEM and an Arabic Noun, and between a French DEF and an Arabic Noun. It is hypothesized that, as it was formulated in part based on TA-French data, the FHC will be shown to hold true for TA-French CS. Similarly, the predictions made by the CAD are expected to hold true for MA-Fr CS. Specifically, it is expected that TA and MA CS with French will differ in regards to switch locations 1, 4, 7 and 8.

3.2 Survey instrument

3.2.1 Use of acceptability judgments

Acceptability judgments, sometimes called grammaticality judgments, have long been used in the study of syntax to determine the sentences that are and are not possible in a given language (Myers 2009a, Featherston 2007). Chomsky's famous sentence, “Colorless green ideas sleep furiously,” (Chomsky 1957, 15), grammatically perfect but completely meaningless without an elaborate context, illustrates the reasons that this type of task is now called an acceptability judgment task. A grammatical sentence may not be acceptable due to semantics, pragmatics, or difficulty in parsing. In the past, judgments used in linguistics were often the author's own judgments, or the author's judgments corroborated by a select group of native speakers of the same language. Any attacks on the conclusions presented from such data could be considered relevant to a different idiolect or dialect, leaving the judgments themselves unassailable in some respects. In recent years, there have been calls for greater validity and a more empirical approach to acceptability (Myers 2009a, Myers 2009b, Featherston 2007). While many caution against assuming that acceptability judgments can alone indicate grammaticality, their value as one way to understand language use is generally recognized (Den Dikken et al. 2007, Fanselow 2007, Haider 2007).

Many in the field of code-switching have been wary of using acceptability judgments due to the fact that code-switching is a marked feature of language use in many communities. It may be looked down on to the point that speakers may "reject as impossible the very sentences they have been recorded using." (Pfaff, 1979, 301). This has led many to use corpora to investigate possible syntactic constraints on code-switching. However, relying strictly on corpora may not provide the full range of options available to a speaker. While Bentahila and Davies

(1983) find corpora useful (most of their data came from recorded conversations) they point out that switches that appear rarely for pragmatic reasons may in fact be entirely admissible to fluent speakers of the languages involved. They also note that all speakers occasionally hesitate or make slips of the tongue that may not be generally acceptable or representative of their language use as a whole. These mistakes are present in the corpora and, without any external way to indicate that these are in fact unusual mistakes, they cannot be discounted.

For these reasons, they supplemented their corpus with elicited acceptability judgments. In using judgments, Bentahila and Davies point out several pitfalls to beware when using this type of data. The first is that accepted sentences can be assumed to be grammatical, though rejected sentences are not necessarily ungrammatical. If the respondent does not find any motivation for the presented switch, the sentence may be rejected. As they note, this may lead researchers to state that entire switch locations are impossible when, in fact, it was only the examples used in elicitation that contained a problem.

3.2.2 Authenticity rating scale

While they are not without flaws, a form of acceptability judgment tasks are used in the current study as an empirical way to examine whether differences exist in code-switching between Moroccan Arabic-French and Tunisian Arabic-French. The chosen variant of acceptability, authenticity, relates to the knowledge that the respondents have about their dialects. The actual scale used by respondents is given in section 3.3 below. A rating scale based on authenticity chosen in order to avoid notions of grammaticality and prescriptivism, which are rarely associated with code-switching, putting the focus strictly on how language may be used within the speech community. The scale originally envisioned by the researcher asked for the likelihood

of hearing each phrase. However, on showing the completed survey to a native speaker, that speaker's understanding of the directions was to rate the likelihood of the given topics. After discussing the true target information with the native speaker, the scale of authenticity was suggested and consequently used in the final survey.

3.2.3 Survey stimuli

60 sentences in French were written by the researcher and verified by a native speaker of French for acceptability. 5 sentences were written for each switch type, for a total of 40 target sentences. 20 of the sentences start in Arabic and switch to French; 20 do the reverse. There is only one switch per sentence. In order to clearly establish for the listener the language before and after each switch, adverbial phrases were added to the beginning or end of the sentence to verify that more than one word was in each language, as seen in the beginning of (60).

(60) *Koula lila, l* enfants mangent les dattes.

'*All night long, the* children eat dates.'

For the switches between an article and a noun, only definite articles were used. This was chosen in part for consistency because the demonstrative determiner must be followed by a definite article in each of the dialects of Arabic. The demonstrative determiners used in sentences involving a switch between a demonstrative and a definite article are all of the type that require a definite article. In the Tunisian version of the survey, the form 'ha' was used exclusively in the survey as the native speaker found it more appropriate. The form 'had' was used exclusively in the Moroccan version of the survey for the same reason.

The remaining 20 sentences used in the survey were created as fillers. These sentences

also contained switches at certain syntactic locations, all outside of the noun phrase, used to distract respondents from the target switch locations. 10 sentences were switched after a preposition and 10 between a subject and verb. For each filler switch type, 5 sentences started in French and ended in Arabic and 5 started in Arabic, ending in French.

After this first stage of materials preparation, a native speaker of each dialect translated each sentence to switch codes at the desired location. Borrowings that have identical or near-identical forms in both languages were avoided at switch locations, such as 't-shirt' which is used in French, Tunisian Arabic and Moroccan Arabic. However, when the only form of a word found appropriate followed the phonology of the borrowing language, such as 'tobis' (bus) for the Moroccan form of French 'autobus,' this was allowed. These sentences were then recorded by an adult male native speaker of each target dialect of Arabic and checked by another native speaker to verify that the sentences were switched at the appropriate locations, both of whom happened to be female, before being uploaded as part of the survey. While the speakers verifying the switches felt that some sentences were awkward due to the switch, they agreed that the sentence fragments before and after the switches were well-formed. The sex of the native speakers making the recordings was purposely selected to be the same in order to avoid any bias that may exist due to sex of the recorded speaker, observed by Lawson and Sachdev in Tunisian Arabic (2000). As a male speaker was the first recruited to make recordings, the other speaker used for recordings was also male. The sex of the native speakers who checked the sentences was purely coincidental. For the Moroccan recording, a female native speaker also translated the sentences before the recordings were made. Due to this fact, the sentences in Moroccan Arabic were also checked by the male native speaker who made the recordings.

3.3 Procedure

One survey was created for each target dialect, making two completely separate versions of the survey. In an effort to achieve continuity in the survey experience, the survey could only be executed once per participant, during one sitting. Each participant first completed a short training session with 2 grammatical sentences in French, 2 grammatical sentences in the dialect, and 2 sentences with completely mixed word order in French and the dialect. During the training session and the main survey, the participants were directed to listen to the recorded sentences and rate each sentence on a 5-point likert scale according to how authentically Tunisian or Moroccan he or she judged the sentence to be. For each survey, only the country-specific adjective appropriate for the dialect in question was used. 1 was considered to be the lowest score and the rating scale, and 5 the highest. The descriptions that accompanied the scale during the training session can be seen in (61) below.

(61) 1 – Très peu tunisienne/marocaine
Not at all Tunisian/Moroccan
2 – Assez peu tunisienne/marocaine
Not very Tunisian/Moroccan
3 – Quelque peu tunisienne/marocaine
Somewhat Tunisian/Moroccan
4 – Assez tunisienne/marocaine
Rather Tunisian/Moroccan
5 – Très tunisienne/marocaine
Very Tunisian/Moroccan

The training session verified that the subjects understood the scale and were comfortable using the online survey website. After the training session was completed, all 60 sentences were given in one sitting. In the main portion of the survey, the respondents were presented with a link titled “Phrase #” with the sentences ordered by appearance in the survey. The link opened a new window in which the sentence would play. Respondents were directed to

listen to each sentence a single time before rating it; however no mechanism was available to prevent a respondent from playing the sentence multiple times. The sentences were presented one per screen and respondents were required to rate a sentence before continuing to the next. Once the next page loaded, it was impossible for respondents to return to the previous page and re-listen to a sentence or change a response. No response times were recorded for analysis. The number of times a sentence was played was also not recorded.

For the main rating portion of the survey, the descriptions in French were no longer given, but participants were asked to continue rating according to the same scale. The goal of this was to have as little text as possible present on the screen as the participants listened to the target sentences. The basic survey directions such as “next” and “quit survey” as well as the message “This question requires a response” were all given in French, but as the language of the survey instrument may affect the responses given about code-switching (Lawson &Sachdev 2000), a minimal amount of written language was desired in order to least affect the respondents’ stated opinions.

After rating the sentences, the participants completed a short language background questionnaire. The full language background questionnaire can be found in Appendix A, which includes questions on the languages spoken by the participant, languages of schooling, current language use, and basic biographical information. These are not used in post-hoc correlations in the current study as described in section 3.4.

The surveys took place in varying settings where participants had access to the internet, but all participants were directed to complete it in a quiet space using a pair of headphones if possible, or a good set of computer speakers if headphones were not available. Participants were directed to not ask the opinion of anyone else, and to not have anyone else in the room if

using speakers instead of headphones in order to eliminate potential bias that may come from reactions of another speaker.

3.4 Subjects

Many of the participants in the study were found through personal contacts of the author. An email announcing the survey was sent to native speakers of Tunisian and Moroccan Arabic who then forwarded the email to other native speakers who were considered by the sender to speak French fluently. In addition, an announcement regarding the study was made through social networking media and further participants were recruited using this method. All responses to the survey were gathered during September and October, 2010.

35 native speakers of Moroccan Arabic and 32 native speakers of Tunisian Arabic completed the survey. The introductory page to the survey explained that participants should be native speakers of the Arabic dialects in question who also consider that they speak French fluently. No test of proficiency was given, though the respondents were all asked to rate their skills in French and Arabic in many different areas as part of the Language Background Questionnaire.

Only responses from participants who completed the Language Background Questionnaire in a meaningful way were used for the survey. This eliminated 2 Moroccan speaking and 3 Tunisian speaking respondents who wrote responses unrelated to the questions asked. Any responses for which the birth year indicated that the speaker was younger than 18 years old were also removed from the data, totaling 2 Moroccan and 1 Tunisian Arabic speakers.

A total of 28 valid response sets (14 male, 14 female) were collected for Tunisian Arabic. The ages ranged from 19 to 49 with a mean age of 27.8 years (mode and median 26). 21

respondents were currently living in Tunisia with 1 living in the Ukraine, 1 in Singapore, 1 in Spain and 4 in the US. All respondents were born in Tunisia and all reported speaking Tunisian Arabic as their first language. One was a simultaneous bilingual, 26 spoke French as their second language, and 1 reported French as his 3rd language, after English. The overall range for learning French is from 4 to 10 with a mean of 7.6.

Data from a total of 31 respondents (16 male, 15 female) was collected for Moroccan Arabic. The ages ranged from 19 to 56 with a mean age of 27.8 years (mode 21, median 23). 27 respondents were currently living in Morocco with 1 living in Canada, 1 in France and 2 in the US. All of the subjects were born in Morocco and reported speaking Moroccan Arabic as their first language. 2 were simultaneous bilinguals, 28 spoke French as their second language, and 1 reported French as his 3rd language. The overall range for learning French is from 0 to 15 with a mean of 5.7 years. Due to the extremely small range of average ratings by subject in the data collected, the variables of current residence and age of acquisition were not correlated with average authenticity ratings. Any differences that do exist are too small to be meaningful.

4 Results

In order to explore the research questions and hypotheses presented in section 3.1 , the results here will be examined first through descriptive statistics, then through statistical tests to identify any significant differences that may exist or interactions between the variables.

Code-switching models make very different predictions about switch types depending on the language of the head of the phrase and the use of code-switching may vary as a result of the language used before or after the switch. For these reasons, the switches here have been analyzed as eight separate types, or levels of the independent variable, even though the grammatical elements on either side of the switch are the same regardless of language.

4.1 Descriptive statistics by variable

One focus of this study is whether the models presented above are accurate across dialects. If any model is, the results across all respondents will reflect this. Table 3 gives the average switch rating by type across all respondents. This table, and all that follows, use only the type number. Examples of all types can be found in section 2.2.2.3. Each switch type begins in the first language indicated and changes to the second at a particular syntactic location. Switch types 1 and 2 occur between a DEM and DEF, 3 and 4 between a DEM and N, 5 and 6 between a N and A, 7 and 8 between a DEF and N.

Table 3 - Average by switch type across all respondents

	Arabic-French				French-Arabic			
	1	3	5	7	2	4	6	8
Mean	2.99	3.00	3.11	2.83	2.46	2.74	3.13	2.78
StDev	1.32	1.36	1.39	1.33	1.35	1.35	1.35	1.41

The most striking feature of the results as a group is the high degree of neutrality; all of the average ratings are near 3 on the scale of 1 to 5, defined as "Somewhat Tunisian/Moroccan." All of the average ratings are within 0.67 points of each other. None of the switch types are rated near 1, "Not at all" authentic, or 5, "Completely" authentic as would be expected for ratings strongly influenced by the acceptability of each switch type. Figure 3 below shows the average authenticity ratings by switch type for all respondents.

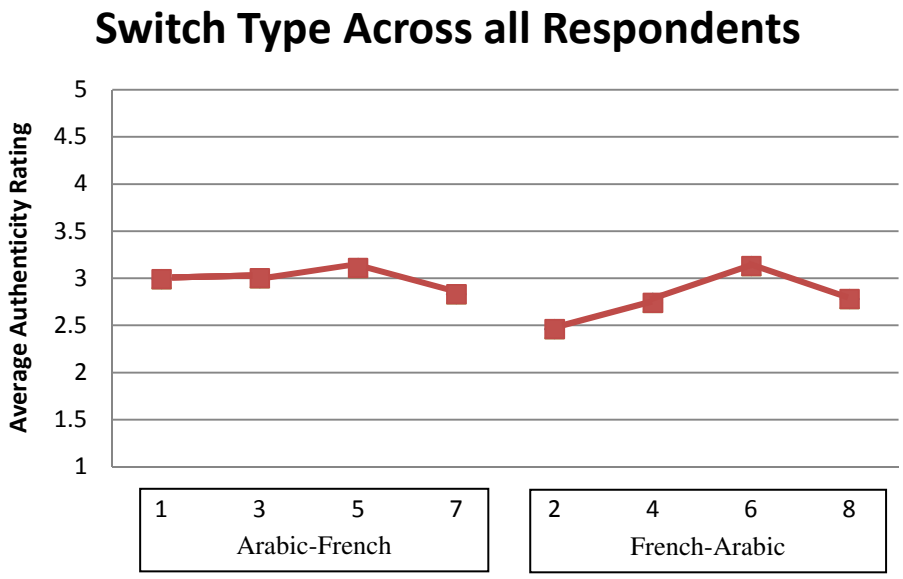


Figure 3: Average authenticity ratings of switch types across all respondents
Each switch type begins in the first language indicated and changes to the second at a particular syntactic location. Switch types 1 and 2 occur between a DEM and DEF, 3 and 4 between a DEM and N, 5 and 6 between a N and A, 7 and 8 between a DEF and N.

It was not expected for all switch types to receive similar ratings, but Figure 3 indicates that they are in fact quite close together. The lowest rating for switch type is found in location 2, from a French Demonstrative Determiner to an Arabic Definite Determiner, with an average rating of 2.46 across all respondents. It bears noting that this is one of the locations in which a code

switch is prohibited by all three of the models discussed above, although its rating of 2.46 is not strongly negative. The highest rated location is type 6, from a French Noun to an Arabic Adjective, with an average authenticity rating of 3.13, followed closely by type 5, a switch from an Arabic Noun to a French Adjective, with an average of 3.11. These two types are permitted in all of the models discussed above. While they do receive the highest ratings, the fact that the average ratings are barely above 3 does not show robust preference. The results across all respondents indicate that switch type does not greatly affect authenticity ratings.

The relatively large standard deviations within switch types for each type indicate that there is a high degree of variability in stated preferences within each switch type. Considering that the averages are near the "neutral" score of 3, the standard deviation of about 1.33 means

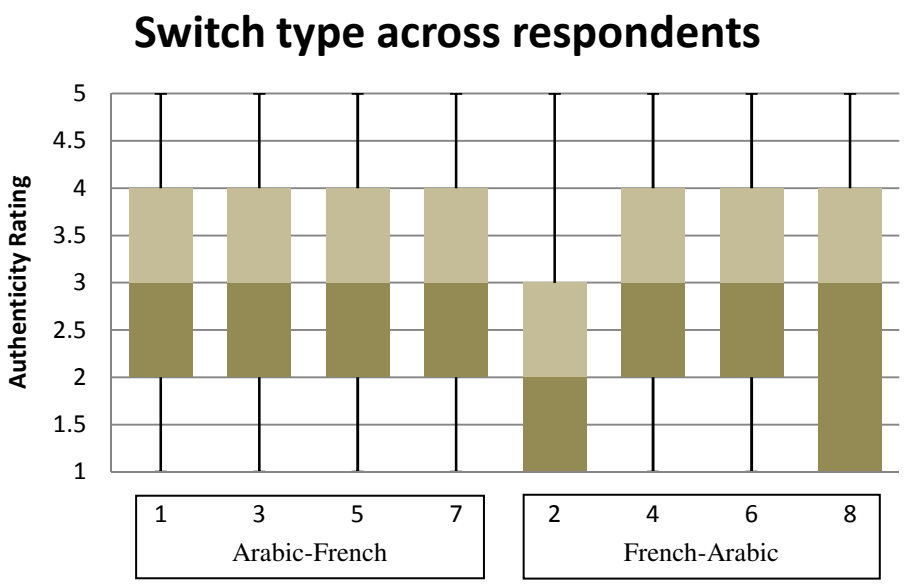


Figure 4: Box plot of switch types across all respondents displaying the median, 1st and 3rd quartiles, and range of the responses by switch types. The median is seen at where the two shaded rectangles meet. The 1st quartile is at the bottom of the darker shaded rectangle. The 3rd quartile is at the top of the ligher shaded rectangle. The range is indicated by the black lines extending out from the quartiles.

that the range of switch type ratings spans the whole given scale, from 1 to 5. This can also be observed graphically in the box plot seen in Figure 4. The whiskers indicate the range of ratings by type, showing that ratings for each switch type span the whole possible range of 1-5. These results suggest that as a full group there is little agreement about what makes a given sentence "Tunisian" or "Moroccan." The medians, indicated in Figure 4 where the shaded rectangles meet, reveal that type 2, from a French DEM to an Arabic noun, has a median rating a full point lower than the other switch types tested. Figure 4 also shows that switch type 8, from a French DEF to an Arabic noun, is more skewed to the lower end of the scale than other types, despite the fact that its median rating is similar to the rest. The box plot shows that the respondents did not all give ratings near the center of the range to each individual stimulus; there are ratings of 1 and 5 for every switch type with no apparent inter-subject agreement. The two departures from 3 are in the direction of lower authenticity, but not to an extent that might be expected for prohibited switch locations.

The lack of clear difference between switch types is not predicted by any of the models discussed. Table 2 shows the permitted types by model and is repeated here as Table 4

Table 4: Permitted and prohibited code-switch types by CS model

	Arabic-French				French-Arabic			
Model	Type 1	Type 3	Type 5	Type 7	Type 2	Type 4	Type 6	Type 8
FHC	N	N	Y	N	N	N	Y	N
CAD	Y	N	Y	Y	N	Y	Y	Y
FHSC	Y	N	Y	Y	N	Y	Y	Y

Each switch type begins in the first language indicated and changes to the second at a particular syntactic location. Switch types 1 and 2 occur between a DEM and DEF, 3 and 4 between a DEM and N, 5 and 6 between a N and A, 7 and 8 between a DEF and N.

Results for switch Types 2 and 3 go against the expectations of all three models while Types 5 and 6 are expected to be permissible. Types 5 and 6 do receive the highest ratings, but the

range of ratings for these types is still the same as all other types. Type 2 receives the lowest rating, but the range of ratings again overlaps with the rest. Type 3 receives the 4th highest ratings of the group, completely unexpected based on the CS models as all prohibit a switch between an Arabic Det and a French Noun. Even so, with an average of 3, described by the rating "Somewhat Tunisian/Moroccan," this cannot be considered evidence for the authenticity of this switch type.

The next question, then, is whether the pattern of neutrality persists when the data is divided by dialect as each group on its own may depart from the overall averages. It is possible that the results by dialect could show a stronger tendency to follow one model over another, which, when combined for all subjects as a whole, might create the appearance of neutrality. In Table 5, the averages are given by dialect.

Table 5 – Switch location by dialect

		Arabic-French				French-Arabic				Overall
		1	3	5	7	2	4	6	8	
Tunisian	Mean	2.73	2.94	3.01	2.54	2.36	2.72	3.00	2.83	2.77
	Standard Deviation	1.30	1.39	1.40	1.30	1.31	1.36	1.34	1.41	
Moroccan	Mean	3.23	3.06	3.20	3.10	2.54	2.76	3.25	2.73	2.98
	Standard Deviation	1.30	1.34	1.38	1.30	1.39	1.34	1.34	1.41	

Through the breakdown by dialect, it can be seen that the Tunisian speakers rate all code-switching types, other than 8, lower than the Moroccan speakers, indicating that Tunisians view the examples of code-switching given in the survey to be slightly less authentic than do the Moroccans. In addition, neither group finds any switch location to be particularly acceptable. The overall average for Tunisians does depart somewhat from 3, although the range of averages is virtually the same as the group as a whole with 0.65 points of difference between the highest

and lowest switch locations. The average for Moroccan respondents is near 3, with a marginally larger range in the averages by type with 0.71 points of difference.

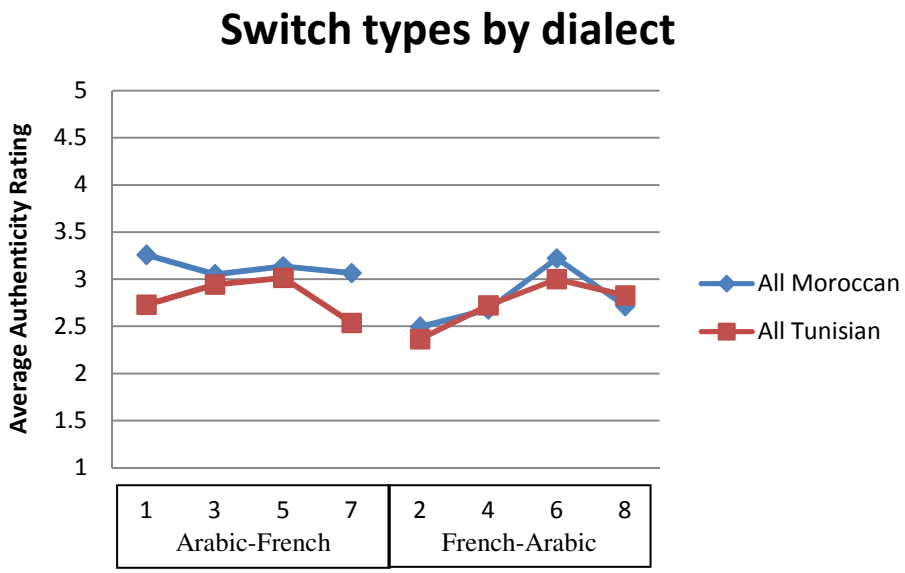


Figure 5: Average authenticity ratings of switch types by dialect

Figure 5 presents graphically the means of the data in Table 2. The Tunisian respondents give more negative responses than the Moroccans with the highest average rating at 3.01, while the Moroccans do show some tendencies to more favorably rate the switch types. The figure makes it plain that for switch type 1 neutrality was indeed heightened by grouping all subjects. The average across all subjects is 2.99 for type 1, while Moroccans rate it 3.23 and Tunisians rate it 2.73. These are, however, minor departures from the original average; neither dialect shows a particularly strong preference. For type 7, the Moroccan average rating remains near 3 while the Tunisian ratings are at 2.54. The overall average was made less neutral in this case as the Tunisian ratings lowered the average. 2.54 is not a strong rating of inauthenticity, but it may indicate a difference between the dialects regarding this switch type.

There is also a difference between the direction of the code switch, Arabic to French or French to Arabic, that can be observed in the data. The range of the average Moroccan ratings for Arabic-French switches is 0.17, much smaller than the range of averages over all of the switch types. The Moroccan respondents gave more varied averages for French-Arabic switches, with a range of 0.71. If a sentence is started in Moroccan then the syntactic location at which a switch to French is made does not affect the ratings as much as it does for a switch from French to Moroccan. The Tunisian respondents show a different pattern by switch direction; the average Tunisian ratings from Arabic to French have a range of 0.47 and French to Arabic with a range of 0.64. For the Tunisians, the language before and after the switch may make a small difference, but not as much as it does to Moroccan speakers.

The results by dialect largely confirm the pattern of neutrality seen in the overall results, but averages by sex may reveal departures from neutrality this. The average rating by type,

Table 6 – Mean switch type rating by sex

		Arabic-French				French-Arabic				
		1	3	5	7	2	4	6	8	Overall
Females	Mean	3.21	3.18	3.34	2.99	2.61	3.05	3.32	2.91	3.073
	Standard Deviation	1.30	1.39	1.37	1.35	1.36	1.30	1.37	1.38	
Males	Mean	2.77	2.83	2.89	2.67	2.46	2.44	2.94	2.65	2.679
	Standard Deviation	1.32	1.31	1.38	1.29	1.35	1.32	1.30	1.43	

separated by sex, can be seen in Table 6. The female average rating is noticeably higher than that of the male respondents, but again there is a very small range of averages, 0.73 points for

females and 0.48 for men. The statistical analysis of this difference can be found in section 4.2 below. Figure 6 shows that the females tend toward preferring Arabic-French switches but are less so of French-Arabic switches. Like

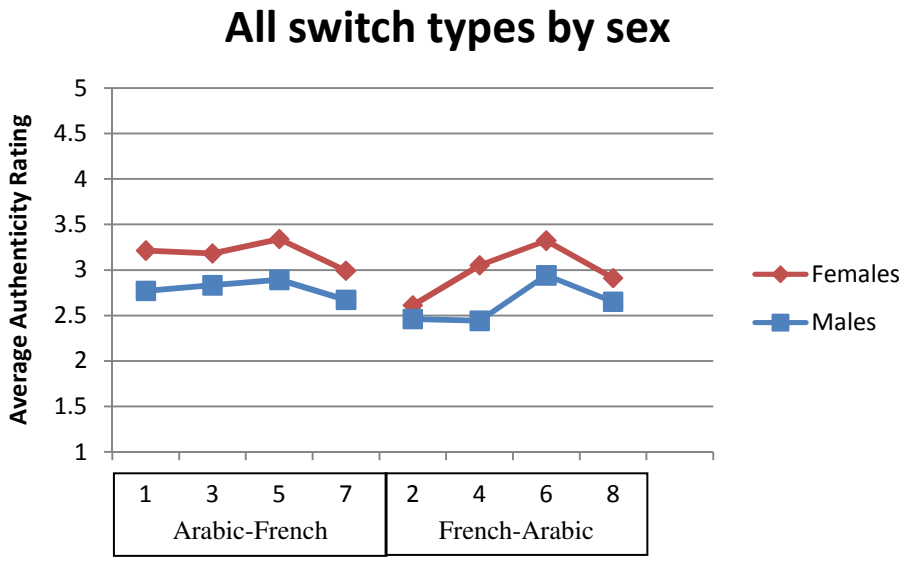


Figure 6: Average authenticity ratings of switch types by sex

the group as a whole, females show a modest peak for switches 5 and 6, between a noun and a post-nominal adjective, but these averages are only 3.34 and 3.32 respectively. They cannot be seen as fully preferring the switch location. When the males diverge from 3, it is negatively, but their ratings do not seem to correspond in any wayto the predictions made by the code-switching models as they rate switch type 2 very similarly to type 4. Both of these switch types are prohibited by the FHC, but types 1, 3, 7 and 8 also are and receive higher ratings.

The last way to divide the group is by both dialect and sex. This is done in Table 7. This again reveals greater individuation between the groups, but does not match any model. The Tunisian Females keep the average rating of 3, with a range of 0.62 between the averages by type. Tunisian Males have the lowest average rating overall, 2.64, with a range of 0.78,

indicating a greater range of averages than either dialect or sex alone. The Moroccan Females have the highest average at 3.30 and a range of 0.9 indicating the greatest amount of distinction

Table 7 – Mean switch type rating by dialect and sex									
	Arabic-French				French-Arabic				
	1	3	5	7	2	4	6	8	Overall
Tunisian F	2.93	3.13	3.19	2.59	2.64	3.26	3.13	3.16	3.00
Tunisian M	2.53	2.76	2.84	2.49	2.09	2.19	2.87	2.50	2.64
Moroccan F	3.47	3.23	3.48	3.36	2.59	2.85	3.49	2.68	3.30
Moroccan M	2.99	2.90	2.93	2.84	2.50	2.66	3.00	2.79	2.89

between switch types of any group. The Moroccan Males again have a lower average, 2.89, and a range of 0.5, revealing the lowest level of distinction between types. The Moroccan females may present the greatest possibility to examine the models presented above as they show the

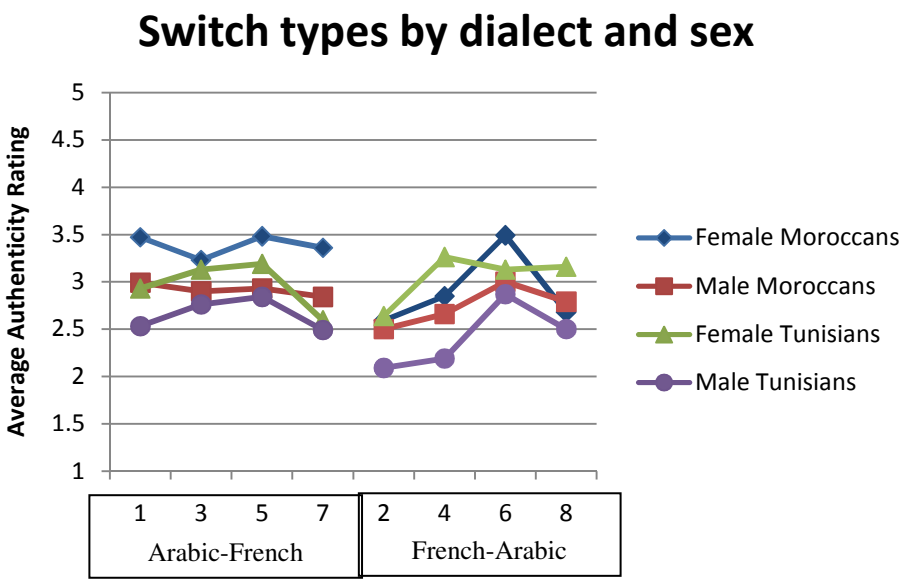


Figure 7: Average authenticity ratings of switch types by dialect and sex

greatest extremes within a single group. However, when viewed as in Figure 7, any such findings seem unlikely. The range of averages for Moroccan Females may be the largest, but they are often the group that most prefers each switch type, not the most sensitive to differences. Moroccan Females give the highest ratings for 5 of the 8 switch types and Tunisian Males give the lowest for all 8 types.When these groups are separated from the whole, a greater degree of divergence from 3, “Somewhat Tunisian/Moroccan” can be seen. The deviations from the overall average for each group is in a single direction, showing one overall tendency within that group; the Moroccan females give higher ratings than the Tunisian males, meaning that they consider all types of code-switching more authentic.

If one group gave both higher and lower ratings than the rest, or if each group gave ratings other than 3 more often, then it would be possible to say they the groups have different preferences by type. However, no group appears to be clearly more or less sensitive to switch type than the others. While the observed preferences indicate varied approval of all code-switching, there are no strong departures from the original averages. Evidently the type does not make a large difference; each group approves or disapproves to a different degree regardless of syntactic elements on each side of the switch. Tests of statistical significance are required to show whether the small observed differences by location are significant. These are reported in section 4.2

4.2 Descriptive statistics by switch type and sentence

None of the above comparisons reveal robust or definitive results. When split into groups, there are differences between male and female respondents, and between dialects, but the rating for each location remains essentially neutral. It is possible that the averages by type

were affected by the individual stimuli used; if a word was used that was not deemed appropriate, this could have decreased authenticity ratings for the sentence and, by extension, lowered the average for an otherwise acceptable switch location. The dialects have different lexicon, and different stimuli for that reason. The averages by type and sentence are therefore examined as a group and by dialect. Individual sentences will be examined as possibly affecting the overall average of the switch type if they differ notably from the other sentences of the same switch type. This difference is operationalized here as more than 0.5 standard deviations from the mean of the switch type. All sentence averages are less than 1 standard deviation from the mean by type.

The first switch type, between an Arabic DEM and a French DET, has the second highest level of divergence in average ratings between the dialects. It can be observed in Table 8 that within this type there is a moderate degree of variability by sentence, though all receive ratings that range from 1-5 and fairly neutral averages.

Table 8: Switch Type 1, Arabic DEM to French DEF

Sentence	14	23	24	36	52	Overall
All	3.37	3.58	2.66	2.66	2.66	2.99
	1.40	1.05	1.25	1.36	1.24	1.32
Moroccans	3.58	3.68	3.06	3.13	2.65	3.23
	1.36	0.94	1.34	1.34	1.28	1.30
Tunisians	3.21	3.52	2.31	2.24	2.76	2.73
	1.45	1.18	1.11	1.30	1.27	1.30

The average rating given by Moroccans for this switch type is higher than that given by Tunisians, but there is some variation by sentence with Tunisians rating sentence 52 above Moroccans. The Tunisian average for sentence 23 is 3.52 with standard deviation of 1.18. This is also the highest rated sentence of this type by the Moroccans, with an average rating of 3.68

and a standard deviation of .94. Compared to the rest of the sentences in this survey, 0.94 is a low standard deviation. Sentence 23 is seen in (62) below. All examples are in Moroccan Arabic-French for consistency. All sentences used in the survey can be found in Appendix B.

(62) *L-bent lhih dart had* les gâteaux il y a deux jours.
'The girl over there made these (the) cakes two days ago.'

There is a difference between this sentence and the others in type 1 is the syntactic and semantic structure on each side of the switch. In sentence 23, most of the main clause is in Arabic with only a DP in French, followed by a temporal adverbial phrase. This form differs from the others of this type as all the others begin with an adverbial phrase in Arabic with the only Arabic element of the main clause the DEM. It may be that Tunisian speakers prefer a switch that keeps as much of the main clause in one sentence as possible. This switch type was also the highest rated for Moroccan speakers, although the average for sentence 23 of 3.68 is much closer to the average Moroccan authenticity rating for type 1. If it is the form that has made a difference for the Moroccan speakers, they may prefer to keep a main clause in one language, but not as much as the Tunisians. It is also possible that the higher ratings here are due to the lexical items used or common collocations and not related to the language of the main clause. Other sentences that diverge from their switch type rating will help answer this question.

Switch type 3, between and Arabic DEM and a French Noun, received very uniform

Table 9: Switch Type 3, Arabic DEM to French Noun						
Sentence	6	12	19	28	53	Type 3
All	3.05	3.53	3.05	2.95	2.44	3.00
	1.25	1.41	1.25	1.31	1.39	1.36
Moroccans	3.16	3.26	3.26	2.87	2.74	3.06
	1.29	1.44	1.24	1.20	1.48	1.34
Tunisians	2.93	3.82	2.82	3.04	2.11	2.94
	1.21	1.33	1.25	1.43	1.23	1.39

ratings by the Moroccans across sentences with a range of 0.52 between the sentence averages. This data is shown in Table 9. The Tunisian data is remarkably different; here there are two sentences with average standard deviations more than 0.5 from the mean by type. More intriguing still, these are in opposite directions; 12 has an average rating of 3.82, higher than the average by type and the average for 53 is 2.11, lower than the average. Here all 5 sentences begin with an adverbial phrase in Arabic; this was the difference seen in sentence 23 of type 1, but cannot account for the wide discrepancy between sentences 12 and 53. Sentence 12 is seen in (63) and 53 in (64)

(63) *Bhall dima had* train arrive en retard.
'*As always, this* train is arriving late.'

(64) *Mouakharren had* chiens aboient toute la nuit.
'*Lately these* dogs bark all night long.'

There are no clear syntactic or semantic reasons for these switch types to receive contradictory ratings. It may be that it is a case of lexical choice, collocation or some other unknown factor that motivated these ratings. The average for this type, however, would remain similar without these sentences as their ratings diverge from the average the same amount in either direction.

In type 5, between an Arabic Noun and a French Adjective there is again a wide amount of variation, this time among the Moroccan respondents, seen in Table 10.

Table 10: Switch Type 5, Arabic Noun to French Adjective

Sentence	13	17	33	47	56	Type 5
All	2.66	2.54	3.14	4.10	3.10	3.11
	1.47	1.25	1.40	1.05	1.23	1.39
Moroccans	2.42	2.55	3.65	4.03	3.32	3.20
	1.41	1.34	1.33	1.05	1.05	1.38
Tunisians	2.93	2.54	2.57	4.18	2.86	3.01
	1.51	1.17	1.26	1.06	1.38	1.40

Sentence 13 is rated low by the Moroccans at 2.42. The main clause of sentence 13, seen in (65), is mainly in French, as are the main clauses of sentences 17 and 33.

(65) *Bla khatrhoum talaba* fatigués preparent leurs examens finaux.
'*Unwillingly, the* tired *students* study for their final exams.'

It may be that it is another case of lexical choice or collocation that earns this sentence low ratings. The adjective 'fatigués' does not show gender in French, making a mismatch between the noun and adjective's gender impossible as a cause for a lower rating. It is interesting that sentence 17 is also rated low, though remains within 0.5 standard deviations of the type mean. Both sentences 13 and 17 contain switches in the middle of the main clause, but as sentence 33 does the same and receives a much higher rating of 3.65, this cannot be the cause of lower ratings.

At the other end of the spectrum for switch type 5 is sentence 47. This is the highest rated sentence of the entire study and the only to receive no ratings of 1 by Tunisians and just one by Moroccans. Sentence 47 is given in (66).

(66) *Hia Kat fddel sbabbt* noires en général.
'*She prefers* black *shoes* in general.'

It can be seen that it is only the adjective 'black' and an adverbial phrase that are in French here. This limited use of French is similar to sentence 23 in Type 1, but does not always garner higher ratings. It could be that both switch type and frequency of vocabulary and collocations affect the ratings, making it difficult to find clear patterns in the ratings when only the switch type is controlled for in the current experiment.

Switch type 7, between an Arabic DEF and a French Noun did not receive any average ratings by sentence that are far from the mean by type. Aside from the switch type, there is no syntactic structure found within this type that is not found in the other types. This leaves no

clear reason for the ratings of sentences across this type to be more homogeneous than others unless due to collocation or lexical items.

Table 11: Switch Type 7, Arabic DEF to French Noun						
Sentence	8	9	26	41	49	Type 7
All	2.59	2.80	2.76	3.22	2.76	2.83
	1.37	1.21	1.47	1.27	1.26	1.33
Moroccans	3.16	2.87	3.13	3.48	2.81	3.10
	1.39	1.31	1.38	1.09	1.28	1.30
Tunisians	1.96	2.71	2.36	2.93	2.71	2.54
	1.04	1.12	1.47	1.41	1.27	1.30

Sentence 8 receives the lowest rating by Tunisians out of all the target stimuli in the survey, but remains near the mean for this type.

Switch type 2, from a French DEM to an Arabic DET, receives the lowest rating for type overall and for each dialect. The range of average sentence ratings is 0.65 for Moroccans, with

Table 12: Switch Type 2, French DEM to Arabic DET						
Sentence	4	29	42	46	58	Type 2
All	2.78	2.58	2.08	2.42	2.42	2.46
	1.49	1.33	1.26	1.30	1.32	1.35
Moroccans	2.71	2.65	2.06	2.58	2.71	2.54
	1.57	1.33	1.29	1.41	1.30	1.39
Tunisians	2.86	2.50	2.11	2.25	2.11	2.36
	1.41	1.35	1.26	1.17	1.29	1.31

four of the five sentence averages within 0.13 points of each other. Sentence 42 is rated 2.06, barely above “Not very” authentic. As the others pattern tightly together, this sentence merits further attention and is given here in (67).

(67) Tu aimerais lire ce *l-ktab fle bhar*.
‘You would like to read this *(the) book on the beach*.’

Again the pattern of a switch before a determiner phrase followed by an adverbial phrase occurs, but it also does in all four other sentences of this switch type. Once more it does not seem to be simple syntax that is motivating the ratings of the sentences in this type. The Tunisian averages by sentence for this type are all well within 0.5 standard deviations of the type mean.

Switch type 4, from a French DEM to an Arabic Noun is the first in the study for which the switch averages are constant across the type with none more than 0.5 standard deviations from the mean and none with an average rating notably different from the rest for either dialect.

Table 13: Switch Type 4, French DEM to Arabic Noun						
Sentence	3	11	20	44	55	Type 4
All	2.71	2.90	2.37	2.73	2.98	2.74
	1.35	1.35	1.23	1.40	1.36	1.35
Moroccans	2.74	3.00	2.26	2.58	3.19	2.76
	1.29	1.37	1.26	1.39	1.28	1.34
Tunisians	2.68	2.79	2.50	2.89	2.75	2.72
	1.44	1.34	1.20	1.42	1.43	1.36

The range of sentence averages for Moroccans is rather high, 0.93, but none of the sentences can be picked out as unduly influencing this average. The sentences all are rated in a similar way by the Tunisians, with a range of just 0.49 between their averages.

Table 14: Switch Type 6, French Noun to Arabic Adjective						
Sentence	5	18	34	40	50	Type 6
All	2.81	3.15	3.41	2.76	3.49	3.13
	1.33	1.24	1.23	1.39	1.39	1.35
Moroccans	3.19	3.23	3.39	2.97	3.42	3.25
	1.42	1.18	1.20	1.49	1.43	1.34
Tunisians	2.39	3.07	3.43	2.54	3.57	3.00
	1.10	1.33	1.29	1.26	1.37	1.34

Switch type 6, from a French Noun to an Arabic Adjective, is again fairly consistent, as in Table 14. Table 14 shows that the range for Moroccans by sentence is 0.45, although the range for Tunisians is much broader at 1.18. However, the average rating for each of the sentences by the Tunisian group remains within 0.5 standard deviations from the mean by type, with 2 sentences lower, two higher, and one at the average by type. This makes it hard to pick out any particular sentence as overly influencing the average created by the rest.

The sentences for the final switch type, type 8, from a French DET to an Arabic Noun, again pattern relatively closely together. The range of the average ratings by sentence for

Table 15: Switch Type 8, French DET to Arabic Noun						
Sentence	16	22	27	30	60	Type 8
All	2.46	2.56	2.88	2.90	3.10	2.78
	1.32	1.36	1.58	1.40	1.34	1.41
Moroccans	2.74	2.35	3.03	2.68	2.87	2.75
	1.41	1.28	1.58	1.42	1.36	1.41
Tunisians	2.14	2.79	2.71	3.14	3.36	2.83
	1.15	1.42	1.58	1.35	1.28	1.41

Moroccans is 0.68 and 1.22 for the Tunisians. The Moroccans clearly agree more than do the Tunisians, who show a large range in the averages. However, this range is again spread evenly with no obvious outliers.

It was noted above that the averages by type pattern more tightly for switches from Arabic to French than they do in the reverse. Examining the switch types by sentence reveals another phenomenon, which may be related. The switch types that begin in Arabic and change to French have more sentences that are rated far from the mean by type. The switch types that begin in French and change to Arabic may display wide ranges between sentence averages, but none stand out particularly from the rest. This confirms that the language before and after the

code switch makes a difference to the native speakers, though why this distinction exists is not clear.

4.3 Statistical tests by variable

Any difference between switch types is virtually meaningless due to the fact that the raw differences between average ratings are extremely small, the ranges all span the entire possible range from 1 to 5 and the standard deviations overlap. However, the large amount of data collected, 40 observations for each of 59 respondents, or 2360 total observations, makes it possible to observe whether small differences exist that may be impossible to identify in the overall picture. The assumption of normalcy is not met for the distribution of the data. It is generally accepted that this does not affect the outcomes of an ANOVA with conservative corrections, but in order to avoid any possible doubt a Friedman Test for related samples was performed for switch type. The Friedman test shows that there is a highly significant difference between switch types, X^2 (7) = 77.54, p<.0001.

Having established that a significant difference exists, a Mixed-Model ANOVA was performed to identify any possible effects of independent variables or interactions between them. The Greenhouse-Geisser correction was used as the assumptions of normalcy and sphericity do not hold for the data set. Despite the matching ranges, there is a statistical difference by switch type, $F_{(7, 1865.96)} = 12.194$, $p < .0001$. A main effect for dialect also exists, $F_{(1, 291)} = 4.793$, $p = .029$, as well as a significant interaction between type and dialect also exists $F_{(7, 1866)} = 2.885$, $p = .007$. This indicates that there is indeed a difference in authenticity ratings for different CS types between the dialects. In practice, this difference is relatively small; it would be impossible to guess a respondent's dialect based on her ratings alone. However, the main effect

for dialect suggests that overall the respondents perceive the switch types as more authentic in Moroccan Arabic than in Tunisian. The interaction between dialect and type indicates that the divergences between dialects on relative authenticity, visible in Table 5 and Figure 3, are significant.

The results show that there is a main effect for sex, $F_{(1, 291)} = 15.771$, $p < .0001$. There is no interaction between sex and switch type because the ratings of the female respondents are always higher than those of the males, $F_{(7, 1866)} = .746$, $p = .622$. When authenticity ratings for CS types are considered by both sex and dialect, a three-way interaction is found, $F_{(7,1866)} = 3.66$, $p = .001$.

Significance may have been reached, but with such a large number of observations it is possible to find very small differences. When a high level of significance is observed in data sets with the same range, it is important to also consider the r^2 values of the tests, which indicate the amount of variation that is due to the given variables. For the mixed-model ANOVA, $r^2 = 0.0254$. The variables investigated here do show significant differences, but the variables account for only 2.54% of the observed variation in the data. This leaves a large amount of room for other variables to impact the switch type ratings.

4.4 Descriptive Statistics for a select group of subjects

Slight tendencies away from 'Somewhat Authentic' do exist in the data, some of which are in the direction suggested by the code-switching models. The large amount of variation in the responses suggests that factors other than switch location influenced respondents' ratings, evidenced by variability within switch types, within responses to each sentence and in the r^2 value of the statistical tests. However, might it be that some respondents were more sensitive

to switch type than others? Switch type 2 receives the lowest rating across all groups and is also disallowed by all the CS models considered here. If these models are correct, the ratings of this switch type reflect their predictions most strongly. The data was examined to identify whether any respondents gave this switch type a particularly low rating, defined as an average of less than 2 across the five sentences. A group of 8 Moroccans and 12 Tunisians meet this requirement. The averages by dialect for this group are seen in Table 16 and Figure 6

Table 16 – Switch location by dialect for subjects with rating less than 2 for Type 2

		Arabic-French				French-Arabic			
		1	3	5	7	2	4	6	8
Tunisian	Mean	2.25	2.60	2.67	2.10	1.47	1.98	2.35	2.07
	Standard Deviation	1.24	1.52	1.45	1.26	0.72	1.14	1.29	1.29
Moroccan	Mean	2.78	2.73	2.73	2.38	1.20	2.03	2.85	1.70
	Standard Deviation	1.31	1.36	1.36	1.27	0.46	1.23	1.44	1.09

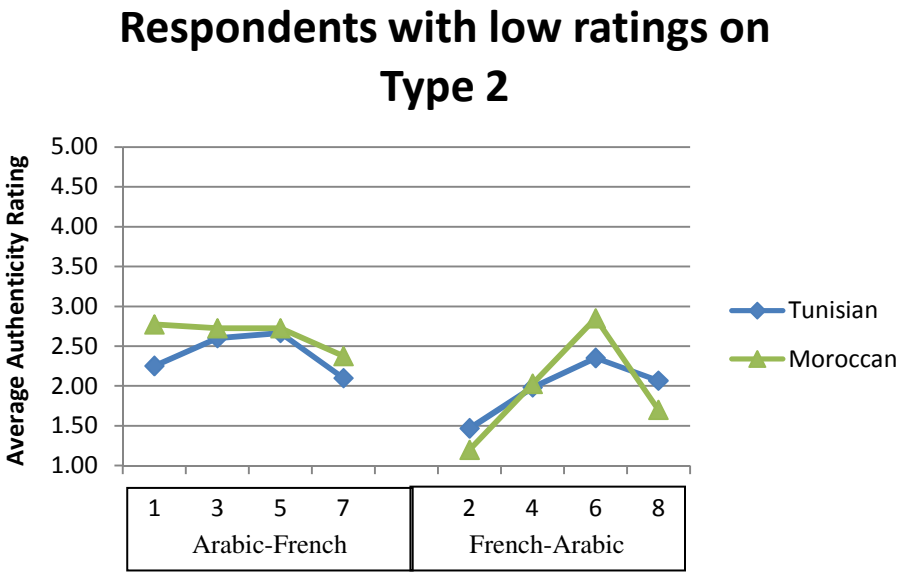


Figure 8: Average authenticity ratings of respondents who rated switch type 2 an average of less than 2.

This group rated all of the switch types lower in general, suggesting an overall dispreference for code-switching. However, a broader range of averages by type is also observed in this group when compared to the averages by dialect. The average ratings by type are quite similar to those given by the full group, with the same overall trends. The group that gave low ratings to type 2 then shows a lower preference for the use of code-switching in general, but the broader ranges indicate that they may also be somewhat more sensitive to the different syntactic locations. Still, none of the models outlined in section 2 above are supported as none of the switch types is viewed as 'authentic' by even one member of this subgroup.

4.5 Discussion

As evidenced by the vast majority of ratings that are near 3, by type and by sentence, the predictions made by the code-switching models discussed above are not strongly supported in the observed results. Either violation of the constraints is not strongly disprefered, or other variables in the sentence overshadowed intuitions respondents have regarding the switch types considered. Both reasons are possible at the same time and imply that the constraints do not operate in an absolute way for code-switching between these dialects and French. The table

Table 17 – Permitted switch types by CS model with average authenticity ratings

	Arabic-French				French-Arabic			
Type #	1	3	5	7	2	4	6	8
CS Model	DEM /DEF	DEM/N	N/A	DEF/N	DEM /DEF	DEM /N	N/A	DEF/N
Average	2.99	3.00	3.11	2.83	2.46	2.74	3.13	2.78
FHC	N	N	Y	N	N	N	Y	N
CAD	Y	N	Y	Y	N	Y	Y	Y
FHSC	Y	N	Y	Y	N	Y	Y	Y

of permitted switch types, originally shown in Table 2, is reprinted here for convenience with the averages by switch type across respondents added in for reference.

None of these models can be upheld by the current results. Above all, the overwhelming result is that none of the types are definitively better or worse than any others. The departures from the rating of 3 are small with wide standard deviations. If these departures can be taken as indications of acceptability of switch types it requires assuming that the syntactic switch types investigated in this study are only one factor considered in rating the sentences. This could be the case, though the addition of this large assumption does not lend credibility to any model discussed here.

Switch types 2 and 3, completely impermissible according to the models, have different average authenticity ratings. Switch type 2 is the lowest average in the data, rated at 2.46 and type 3 is the third highest with an average of 3.00. Neither of these averages indicates that the types are impossible, but the difference between them is further evidence against the accuracy of the code-switching models.

Types 1, 7, 4 and 8, where the models do not align, are rated barely below 3. These switch types are considered impossible according to the FHC, based in part on Tunisian data. If all four of these were rated lower by Tunisian respondents, it may be a sign that their responses trend in the direction of the FHC, suggesting that it may in fact play some role in their stated authenticity ratings. While the Tunisian respondents did rate types 1, 7 and 4 lower than the Moroccans, they rated the type 8 higher and types 3, 5, 2 and 6 lower than the Moroccans as well. It seems that Tunisians prefer fewer of the switch types included in the experiment, but not in a pattern than supports the FHC.

One aspect of all of the results that cannot be ignored is the lack of extreme ratings

across the board. No switch type is strongly seen as authentic, nor is any seen as absolutely inauthentic. In this study, many aspects of the sentences used as stimuli were not controlled. If five sentences had been used in 8 forms, each form of the repeated sentence demonstrating each of the 8 switch types, there may have been a clearer distinction between the types in the ratings given. Yet if the proposed constraints are absolute, then survey respondents should have responded to violations of the constraints even if the rest of the sentence was acceptable.

When asked to rate sentences on 'authenticity' it is possible that syntactic acceptability did play a role because significant differences between the types do exist, despite the fact that they do not account for very much of the observed variation. Perhaps the constraints on CS should be stated as only tendencies. The lowest and highest ratings were given to certain sentences for which the predictions of the CS models agree as being unacceptable and completely acceptable. However, other factors may have captured the attention of respondents, such as accent, amount of each language used per sentence and vocabulary choice. In fact, some survey participants commented on these exact factors. Some asked where the speaker was from, others stated there was "too much French" in the sentences, and yet others pointed out that the vocabulary chosen, perhaps to avoid potential borrowings, was "too standard" and not representative enough of the dialect.

An alternative possibility is that the answer may lie in the way the question was phrased; respondents were asked to give their opinions of the authenticity of the sentences given. In attempting to avoid mention of grammaticality due to the potential disapproval of code-switching, judgments regarding frequency instead of syntactic acceptability may have been triggered. Bentahila and Davies (1983) report a concern that syntactically permissible switches may rarely occur in natural conversation if they are not motivated by semantic or pragmatic

reasons, and thus may be absent from a corpus study. If these switch types occur less frequently in actual use, this could be one source of the tepid authenticity ratings.

5. Conclusion

The results of this study are neither exhaustive nor conclusive. The goal of this study was to explore syntactic constraints in code-switching in the same language pair in different contexts. Within the experiment, four questions were investigated and answered. The first is whether speakers of each dialect indicate the same levels of authenticity for code-switching. It was seen that the Moroccan Arabic speakers rate CS as more authentic than the speakers of Tunisian Arabic. Related to this is the second question of whether speakers of the dialects prefer the same switch types. While the observed preferences are very similar, they are significantly different. The two dialects diverge most notably on the authenticity of switch types 1 and 4 while they generally pattern together.

The third question is whether the sexes state different levels of authenticity. While they do trend together on higher and lower ratings, the females report higher authenticity ratings than the males. Based on these results, females appear to find code-switching more authentic than males. This may be due to the fact that they participate in code-switching more often, as concluded by Lawson and Sachdev (2000) and others.

The fourth and final question was to identify if any model of code-switching put forth based on this language pair could be supported. The answer here is a resounding no. With all of the switch types rated near 3, the exact middle of the rating range, syntactic constraints alone are not responsible for perceived authenticity of a given utterance. This is not to say that syntax plays no role; the significant differences found between ratings of switch types imply that they may on a certain level, but not to a great extent. Many of the observed departures from a rating of 3 are also in the directions suggested by the models. However, even these smaller departures do not clearly pattern with any model seen here.

5.1 Limitations

The major limitation of this study is that it was based on elicited judgments. Despite the need for empirical methods, code-switching is not always considered acceptable, even by those who participate in it regularly. By asking for the authenticity of each sentence, the effect of this negative bias may have been minimized, but cannot be assumed to have been erased completely. Bias against code-switching may be the cause of the lack of reported authenticity. Comparisons with contemporary naturalistic observations are necessary to fully validate the results reported.

The method of carrying out the survey also could have affected the results obtained. As the survey was done via internet, all respondents were required to have internet access as well as access to headphones or speakers. In the countries in question, internet access is not available in every household, making it likely that a high proportion of the survey participants belong to the upper classes. Inclusion of all socio-economic groups, such as by replicating the experiment in-country, would make these results more generalizable. It is also possible that the respondents received some type of feedback from others in the room, whether specifically requested or unintentionally noted, such as by body language if the sentences were heard over computer speakers. Future studies should control for these effects.

5.2 Directions for future research

The results of this study suggest that differences in the judgments for authenticity in code-switching exist between the same language pair in different social contexts. By doing

further work in this vein, it may be possible to better understand how and why individuals and communities make use of two languages at once.

Many avenues for future research are possible. One is to explore the relationship between authenticity, frequency and acceptability. The scale used here is based on authenticity. This may have been interpreted by respondents as relating to frequency, leading them to give switch types that occur less often lower ratings. The question of frequency could be explored through a corpus study while acceptability could be explored through direct elicitation. This could be done by first creating a corpus of current usage of code-switching in the two countries to identify the syntactic environments in which it occurs most often. Once these are identified, it would be possible to directly ask subjects whether they find the possible switch environments to be acceptable based on actual examples. Any syntactic environments not attested in the corpus could be created by minimal manipulations to the corpus data.

One notion not discussed here is identity. Asking a respondent about authenticity brings into question what it means to be 'Moroccan' or 'Tunisian.' This relates to her own identity as part of the group and the ratings she gives may reflect whether she considers code-switching an important part of that identity.

As a significant difference was found in the responses between sexes, the sex of the speaker used to create the stimuli may also have an impact on ratings. Other personal factors of the respondents could also be more carefully controlled for, such as level of education, proficiency level in each language, socio-economic status, other languages spoken, and many more.

The most accurate way to gauge use of code-switching is to gather naturalistic data in-country via recordings of native speakers. This could increase the differences in stated

speaker preferences, or may negate them completely. Either outcome would aid in filling the current gap in the literature. Although more costly and time-consuming, naturalistic data is the only ecologically valid way to be sure of actual syntactic properties of code-switching employed by fluent speakers.

Appendix A – Language Background Questionnaire

For consistency, the Moroccan version of the survey is given here. The Tunisian version is the same other than references to Tunisia and the Tunisian dialect where Morocco and the Moroccan dialect are mentioned here

Sur cette page et les pages suivantes, il y a des questions personnelles. Rappelez qu'il est essentiel que vous répondiez à toutes les questions avec toute franchise. Vos réponses sont anonymes et vous pouvez vous assurez que les résultats ne vous seront pas liés.

1. sexe: M F
2. Où est-ce que vous êtes né(e)?
3. Dans quelle année est-ce que vous êtes né(e)?
4. Dans quelle ville est-ce que vous habitez actuellement?
5. Depuis quand est-ce que vous habitez cette ville? (mois et année)
6. Si vous avez vécu hors du Maroc, écrivez le nom de chaque pays où vous avez habité. Précisez le temps pendant lequel vous y avez habité (ex: 3 mois, 10 ans)
7. Quelle est votre première langue? (Si votre première langue est l'arabe ou un dialecte de celle-ci, veuillez indiquer le pays dans lequel vous l'avez appris.)
8. Quelle était votre deuxième langue? (Si votre deuxième langue est l'arabe ou un dialecte de celle-ci, veuillez indiquer le pays dans lequel vous l'avez appris.)
9. A quel âge est-ce que vous avez commencé à apprendre votre deuxième langue?
10. Quelle était votre troisième langue? (Si votre troisième langue est l'arabe ou un dialecte de celle-ci, veuillez indiquer le pays dans lequel vous l'avez appris.)
11. A quel âge est-ce que vous avez commencé à apprendre votre troisième langue?
12. Ecrivez les autres langues que vous parlez et l'âge à laquelle vous avez commencé de l'apprendre.
13. Quel est votre niveau d'étude?
14. Dans quelle langue est-ce que vous avez fait vos études secondaires?
15. Dans quelle langue est-ce que vous avez fait vos études universitaires?
16. Dans quelle langue est-ce que vous avez fait vos études primaires? (choix : Arabe, Français, Anglais, Autre (indiquez))

    Veuillez répondre aux questions suivantes sur votre usage de français.
17. Est-ce que vous écrivez souvent en français?
18. Est-ce que vous parlez souvent en français?
19. Est-ce que vous écoutez la radio ou la musique en français?
20. Est-ce que vous lisez le journal, les magazines, or les livres en français?
21. Est-ce que vous regardez la télévision ou les films en français?

    Veuillez répondre aux questions suivantes sur votre usage d'arabe tunisien/marocain.
22. Est-ce que vous écrivez souvent en arabe marocain?
23. Est-ce que vous parlez souvent en arabe marocain?
24. Est-ce que vous écoutez la radio ou la musique en arabe marocain?
25. Est-ce que vous lisez le journal, les magazines, or les livres en arabe marocain?
26. Est-ce que vous regardez la télévision ou les films en arabe marocain?

Veuillez répondre aux questions suivantes sur votre usage d'arabe standard.

27. Est-ce que vous écrivez souvent en arabe standard?
28. Est-ce que vous parlez souvent en arabe standard?
29. Est-ce que vous écoutez la radio ou la musique en arabe standard?
30. Est-ce que vous lisez le journal, les magazines, or les livres en arabe standard?
31. Est-ce que vous regardez la télévision ou les films en arabe standard?

Indiquez votre niveau de français dans les domaines suivantes: (1= pas du tout, 3 = moyen, 5 = aussi couramment que ma première langue)

32. Votre capacité de parler le français
33. Votre capacité de comprendre le français
34. Votre capacité d'écrire le français
35. Votre capacité de lire le français
36. Votre prononciation en français
37. Votre grammaire en français
38. Votre capacité globale en français

Indiquez votre niveau de dialecte tunisien/marocain dans les domaines suivantes: (1= pas du tout, 3 = moyen, 5 = aussi couramment que ma première langue)

39. Votre capacité de parler le dialecte marocain
40. Votre capacité de comprendre le dialecte marocain
41. Votre capacité d'écrire le dialecte marocain
42. Votre capacité de lire le dialecte marocain
43. Votre prononciation en dialecte marocain
44. Votre capacité globale en dialecte marocain

Indiquez votre niveau d'arabe standard dans les domaines suivantes: (1= pas du tout, 3 = moyen, 5 = aussi couramment que ma première langue)

45. Votre capacité de parler l'arabe standard
46. Votre capacité de comprendre l'arabe standard
47. Votre capacité d'écrire l'arabe standard
48. Votre capacité de lire l'arabe standard
49. Votre prononciation en arabe standard
50. Votre grammaire arabe standard
51. Votre capacité globale en arabe standard

Appendix B – Survey Stimuli

As for the Language Background Questionnaire, only the Moroccan versions of the sentences are given here in the order in which they were presented in the survey. The spelling given in the Latin alphabet is that used by the consultant native speaker of Moroccan Arabic. The Tunisian versions are the same other than lexical items that differ between the dialects.

1. Cette robe coute peu dans had l7anout.
2. Elle préfère cette chemise avec had srwal.
3. Sa femme a acheté ces karassa f sou9.
4. Pendant le diner, cette l bent m3asba har9at kass-ha.
5. Les enfants ont commencé à aimer les plats m3trine chwia b chwia .
6. Nichane mbe3d had chanson me rappelle mon enfance.
7. Elle prépare le dîner en sa3tayene.
8. Koula lila l enfants mangent les dattes.
9. f 9leb l medina l hommes construisent l'immeuble.
10. Llwalida connait ce poème par cœur.
11. Depuis le mois dernier ce mjer mab9ach tay thherek.
12. Bhall dima had train arrive en retard.
13. Bla khatrhoum talaba fatigues preparent leurs examens finaux.
14. Fel louwel dial dewera had le professeur donne beaucoup de devoirs.
15. Ghan nghsslou lma3an mousskhine be3d le dîner.
16. Les filles lavent le kleb men jdid.
17. Sebt l bent sportive va à la plage.
18. Elle a fait tomber le café skhoune 3la tbla.
19. Wakha dakshi had lampes n'éclairent pas la pièce.
20. La fille sympathique a aidé cet rajjel simana li fatet.
21. Cet homme ne boit pas de lait depuis sghrrou.
22. On a brulé le rroz f l mtbakh.
23. L bent lhih dart had le gâteaux il y a deux jours.
24. Koulla tnine had la femme intelligente travaille à la banque.
25. Samedi soir, le couple mcha l mt3em.
26. L mouaj kay drbou dd l rochers sans cesse.
27. On a regardé souvent les mousslssalates ensemble.
28. Had l aw9at llekhra had etudiants sortent tous les soirs.
29. Elle a fait cette l 7lwa 7llouuwa f sa3a.
30. Depuis une heure le jdd kay 9leb 3la ndadrou.
31. Chrajem f cette salle ne se ferment pas.
32. Koul lila, kan n3ess mbe3d les infos.
33. Asslan Toumoubila rose ne marchent pas.
34. Elle mange plus la nourriture 7arra melli merdat.
35. B7al dima had tallib arrive en retard.
36. Men dima had l'homme n'aime pas les chiens.
37. Ma Femme ghadi tmchi mendaba sa3a.
38. Mmi ghat koun ghdbana 3la ma sœur quand elle la verra.

39. Hia kerkbat lwoueld m3a la porte.
40. La famille évite la plage mouskha f l 3chiate.
41. Houwa Kay 9leb 3la l télécommande depuis ce matin.
42. Tu aimerais lire ce l ktab fle bhar.
43. L khamiate cachent ces fenêtres brisées.
44. Il porte ce fou9ani koula nhar.
45. En été ce chien kay bghi y mchi l 7adi9a.
46. Mon père ne connait pas cette l medina b7al l lma.
47. Hia Kat fddel sbabbt noires en général.
48. En classe, le professeur kay hddar bchwiya.
49. L 9at kay 3ad l souris une fois pour toutes.
50. Il a trouvé une serviette fazgha f lred.
51. Tu aimerais jouer ce jeu avec s7abek.
52. Ghir Rchou9 had l'enfant a cassé le jouet
53. mouakharren had chiens aboient toute la nuit.
54. Une voiture 9lbat moul bchklita.
55. Tous les jours ce tobis kay 7bess f rass derb.
56. Ana Sheft l 3imara moderne en ville.
57. Mnin taykoune 3yane, l 9at griffe le canape.
58. Les rideaux cachent ce l hitt m9cher 7mdoullah/b koul farah.
59. 3la ayi7al, had chikhat tombent souvent.
60. J'ai perdu le ktab f l medrassa.

References

Abbassi, Abdelaziz. 1977. A Sociolinguistic analysis of multilingualism in Morocco. Austin, TX. The University of Texas at Austin dissertation.

Abney, Steven P. 1987. The English noun phrase in its sentential aspect. Cambridge, MA. MIT dissertation.

Bassiouney, Reem. 2009. *Arabic sociolinguistics: Topics in diglossia, gender, identity and politics*. Washington, D.C.: Georgetown Press.

Belazi, Hedi M. 1991. Multilingualism in Tunisia and French/Arabic codeswitching among educated Tunisian bilinguals. Ithaca, NY. Cornell University dissertation.

Belazi, Hedi M; Edward J Rubin & Almeida Jacqueline Toribio. 1994. Code Switching and X-Bar Theory: The Functional Head Constraint. *Linguistic Inquiry* 25. 221-237.

Belfalah, Chedly. 1992. *Syntactic development and retardation across English and Arabic: A contrastive study of normal and hearing-impaired children.* Tunis: Unversité de Tunis I: Faculté de lettres de la Manouba.

Benrabah, Mohamed. 2007. Language-in-education planning in Algeria: Historical development and current issues. *Language Policy* 6. 225-252.

Bentahila, Abdelali & Eirlys, E. Davies. 1983a. The syntax of Arabic-French codeswitching. *Lingua* 59. 301-330.

Bentahila, Abdelali & Eirlys, E. Davies. 1995. Patterns of code-switching and patterns of language contact. *Lingua* 96. 75-93.

Bentahila, Abdelali. 1983a. Motivations for codeswitching among Arabic-French bilinguals in Morocco. *Language and Communication,* 33. 233-243.

Bentahila, Abdelali. 1983b. *Language attitudes among Arabic-French bilinguals in Morocco*. Clevedon: Multilingual Matters.

Bernstein, Judy B. 2001. The DP hypothesis: Identifying clausal properties in the nominal domain. In Mark Baltin and Chris Collins (Eds.) *The handbook of contemporary syntactic theory*. Malden, MA: Blackwell Publishers, 536-561.

Bhatt, Rakesh Mohan. 1997. Code-switching, constraints and optimal grammars. *Lingua* 102(4). 223-251.

Blom, Jan-Petter & John J. Gumperz. 1971. Social meaning in linguistic structures. In J. J. Gumperz. & D. Hymes. (eds.). Directions in Sociolinguistics. N.Y.: Holt, Rinehart and Winston.

Bouchard, Denis. 1998. The Distribution and interpretation of adjectives in French: A consequence of Bare Phrase Structure. *Probus* 10. 139-183.

Brustad, Kristen E. 2000. *The Syntax of spoken Arabic: A comparative study of Moroccan, Egyptian, Syrian and Kuwaiti dialects.* Washington, D.C.: Georgetown University Press.

Bullock, Barbara and Almeida Jacqueline Toribio eds. 2009. *The Cambridge handbook of linguistic code-switching*. New York: Cambridge University Press.

Chan, Brian. 2003. *Aspects of the syntax, the pragmatics, and the production of code-switching*. New York: Peter Lang Publishing.

Chomsky, Noam. 1957. *Syntactic Structures*. The Hague/Paris: Mouton.

Chomsky, Noam. 1993. A minimalist program for linguistic theory. In *The view from Building 20: Essays in linguistics in honor of Sylvain Bromberger*, ed. Kenneth Hale and Samuel Jay Keyser, 1-52. Cambridge, MA: MIT Press.

Clyne, Michael. 1987. Constraints on code-switching: How universal are they? *Linguistics* 25. 739-64.

Den Dikken, Marcel, Judy B. Bernstein, Christina Tortora & Raffaella Zanuttini. 2007. Data and grammar: Means and individuals. *Theoretical Linguistics* 33 (3). 335-352.

DiSciullo, Anne-Marie, Pieter Muysken and Rajendra Singh. 1986. Government and Code-mixing. *Journal of Linguistics* 221. 1-24.

Ennaji, Moha. 1991. Aspects of multilingualism in the Maghreb. *International Journal of the Sociology of Language* 87. 7-25

Ennaji, Moha. 2002. Language contact, Arabization policy and education in Morocco. In Aleya Rouchdy Ed, *Language contact and language conflict in Arabic: Variations on a sociolinguistic theme*, London: Routledge Curzon, 70-88.

Espinosa, Aurelio M. 1911. Studies in New Mexican Spanish, Part II: Morphology. *Revue de dialectologie Romane* 3. 251-286.

Evans, Nicholas and Stephen C. Levinson. 2009. The myth of language universals: Language diversity and its importance for cognitive science. *Behavioral and Brain Science*. 32. 429-492.

Fanselow, Gisbert. 2007. Carrots – perfect as vegetables, but please not as a main dish. *Theoretical Linguistics* 33 (3). 353-367.

Featherston, Sam. 2007. Data in generative grammar: The stick and the carrot. *Theoretical Linguistics* 33 (3). 269-318.

Ferguson, Charles 1959. Diglossia. *Word* 15 325-340.

Gallagher, Charles F. 1964. North African problems and prospects: Language and identity. In *Language Problems of Developing Nations,* J. Fishman, C.A. Ferguson, and J. Das Gupta (eds.), 129-150. New York: Wiley.

Gingras, Rosario. 1974. Problems in the description of Spanish/English intrasentential code-switching. In Southwest areal linguistics, ed. Garland D. Bills, 167-174. San Diego, CA: University of California Institute for Cultural Pluralism.

Gumperz. John J. & Eduardo Hernandez-Chavez. 1971. Bilingualism, bidialectalism and classroom interaction. In Cazden. E. el al. (eds.). *The Functions of language in the classroom*. New York: Teachers College Press.

Haegeman, Liliane. 1991. *Government and binding theory*. Malden, MA: Blackwell Publishing.

Haider, Hubert. 2007. As a matter of facts: Comments on Featherston's sticks and carrots. *Theoretical Linguistics* 33 (3). 381-394.

Harrel, Richard S. 1962. *A short reference grammar of Moroccan Arabic.* Washington: Georgetown University Press.

Holes, Clive. 2004. *Modern Arabic: Structures, functions and varieties*. Washington: Georgetown University Press.

Ibrahim, Muhammed H. 1986. Standard and Prestige Language: A Problem in Arabic Sociolinguistics. *Anthropological Linguistics* 28. 115-126.

Joshi, Aravind. 1985. Processing of sentences with intrasentential code switching. In Dowty et. al Eds *Natural Language Parsing: Psychological, computational and theoretical persepectives.* Cambridge: Cambridge University Press.

Kachru, Braj B. 1977. Linguistic schizophrenia and language census: A note on the Indian situation. *Linguistics* 186. 17-32.

Labov, William. 1991. *Sociolinguistic Patterns*. Philadelphia: University of Pennsylvania Press.

Lahlou, Moncef. 1991. A morpho-syntactic study of code-switching between Moroccan Arabic and French. Austin, TX. The University of Texas at Austin dissertation.

Lawson, Sarah & Itesh Sachdev. 2000. Codeswitching in Tunisia: Attitudinal and behavioral dimensions. *Journal of Pragmatics* 32. 1343-1361.

MacSwan, Jeff. 2009. Generative approaches to code-switching. In Barbara Bullock and Almeida Jacqueline Toribio (eds.) 2009. *The Cambridge handbook of linguistic code-switching*. 309-335. New York: Cambridge University Press.

Mahmoud, Youssef. 1986. Arabic after diglossia. In Charles Albert Ferguson and Joshua A Fishman (eds) *The Fergusonian impact: In honor of Charles A Ferguson on his 60th birthday* 1. 239-252. Berlin: Mouton de Gruyter.

Mahootian, Shahrzad & Beatrice Santorini. 1996. Code Switching and the complement/adjunct distinction. *Linguistic Inquiry* 27. 464-479.

Mahootian, Shahrzad. 1993. A null theory of code-switching. Evanston, IL. Northwestern University dissertation.

Milroy, James. 1981. *Regional Accents of English: Belfast*. Belfast: Blackstaff.

Morocco. 2004. Recensement général de la population et de l'habitat 2004. http://www.hcp.ma/Profil.aspx 9 Oct 2010.

Muysken, Pieter. 2000. *Bilingual speech: A typology of code-mixing.* New York: Cambridge University Press.

Myers, James. 2009a. Syntactic judgment experiements. *Language & Linguistics Compass, 3* (1), 406-423.

Myers, James. 2009b. The design and analysis of small-scale syntactic judgment experiments.

Lingua 119. 425-444.

Myers-Scotton, Carol and Janice Jake. 2009. A universal model of code-switching and bilingual language processing and production. In Barbara Bullock and Jacqueline Toribio eds. The hanbook of linguistic code-switching. 336-357. Cambridge: Cambridge University Press

Myers-Scotton, Carol. 1993. *Duelling languages; Grammatical structure in code-switching.* Oxford, England: Oxford University Press.

Myers-Scotton, Carol. 2010. Short summaries of the Uniform Structure Principle, the MLF Model and the 4-M Model. http://www.myers-scotton.com/short_summaries.htm 17 Sept 2010.

Pfaff, Carol W. 1979. Constraints on language mixing: Intrasentential code-switching and borrowing. *Language* 55 (2). 291-318.

Poplack, Shana. 1980. Sometimes I'll start a sentence in Spanish y termino en español: Toward a typology of code-switching. *Linguistics* 18. 581-618.

Redouane, Rabia. 2005. Linguistic Constraints on Codeswitching and Codemixing of Bilingual Moroccan Arabic-French Speakers in Canada. In Cohen et. al (eds.) *Proceedings of the 4th International Symposium on Bilingualism*, 1921-1933. Somerville, MA: Cascadilla Press.

Rickford, J. 1975. Carrying the new wave into syntax: the case of Black English BIN. In Ralph W. Fasold & Roger W. Shuy (eds.) *Analyzing variation in language: Papers from the second colloquium on New Ways of Analyzing Variation,* 162- 83. Washington, DC: Georgetown University Press.

Sankoff, David and Shana Poplack. 1981. A formal grammar for code-switching. *Papers in Linguistics* 14. 3-45.

Sayahi, Lotfi. 2007. Diglossia and contact-induced language change. *International Journal of Multilingualism* 41. 38-51.

Scipione, Ruth and Lotfi Sayahi. 2005. Consonantal Variation of Spanish in Northern Morocco. In Lotfi Sayahi and Maurice Westmoreland (eds.), *Selected Proceedings of the Second Workshop on Spanish Sociolinguistics*, 127-132. Somerville, MA: Cascadilla Proceedings Project.

Sirles, Craig A. 1999. Politics and Arabization: The evolution of postindependence in North Africa. *International Journal of the Sociology of Language* 137. 115-129.

Timm, L. A. 1975. Spanish-English Code-switching: El porque y how-Not-To. *Romance Philology* 284. 473-82.

Versteegh, Kees. 1997. *The Arabic Language.* Edinburgh: Edinburgh University Press.

Walters, Keith. 2003. Fergie's Prescience: The changing nature of diglossia in Tunisia. *International Journal oc the Sociology of Language* 163. 77-109.

Weinreich, Uriel. 1953. *Languages in Contact*. The Hague: Mouton.

Wodak, Ruth & Gertrude Benke. 1997. Gender as a Sociolinguistic Variable: New Perspectives on Variation Studies. In Florian Coulmas (ed.) *The Handbook of sociolinguistics.* 127-150. Malden, MA: Blackwell Publishing.